Table of Contents

Table of Contents

Data Storage & Transfer

Storage Components

File Systems

Pleiades Home Filesystem

DRAFT

This article is being reviewed for completeness and technical accuracy.

The home file system on Pleiades (/u/username) is an SGI NEXIS 9000 filesystem. It is NFS-mounted on all of the Pleiades front-ends, bridge nodes and compute nodes.

Once a user is granted an account on Pleiades, the home directory is set up automatically during his/her first login.

Quota and Policy

Disk space quota limits are enforced on the home filesystem. By default, the soft limit is 8GB and the hard limit is 10GB. There are no inode limits on the home filesystem.

To check your quota and usage on your home filesystem, do:

```
%quota -v
Disk quotas for user username (uid xxxx):
     Filesystem  blocks   quota   limit   grace   files   quota   limit   grace
saturn-ib1-0:/mnt/home2
                7380152  8000000 40000000          190950       0       0
```

The quota policy for NAS states that if you exceed the soft quota, an email will be sent to inform you of your current usage and how much of your grace period remains. It is expected that a user will occasionally exceed their soft limit as needed, however after 14 days, users who are still over their soft limit will have their batch queue access to Pleiades disabled. If you believe that you have a long-term need for higher quota limits, you should send an email justification to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager, Bill Thigpen, for approval.

The quota policy for NAS can be found here.

Backup Policy

Files on the home filesystem are backed up daily.

Pleiades Lustre Filesystems

Pleiades has several Lustre filesystems (/nobackupp[10-60]) that provide a total of about 3 PB of storage and serve thousands of cores. These filesystems are managed under Lustre software version 1.8.2.

Lustre filesystem configurations are summarized at the end of this article.

Which /nobackup should I use?

Once you are granted an account on Pleiades, you will be assigned to use one of the Lustre filesystems. You can find out which Lustre filesystem you have been assigned to by doing the following:

```
pfe1% ls -l /nobackup/your_username
lrwxrwxrwx 1 root root 19 Feb 23  2010 /nobackup/username -> /nobackupp30/username
```

In the above example, the user is assigned to /nobackupp30 and a symlink is created to point the user's default /nobackup to /nobackupp30.

TIP: Each Pleiades Lustre filesystem is shared among many users. To get good I/O performance for your applications and avoid impeding I/O operations of other users, read the articles: Lustre Basics and Lustre Best Practices.

Default Quota and Policy on /nobackup

Disk space and inodes quotas are enforced on the /nobackup filesystems. The default soft and hard limits for inodes are 75,000 and 100,000, respectively. Those for the disk space are 200GB and 400GB, respectively. To check your disk space and inodes usage and quota on your /nobackup, use the *lfs* command and type the following:

```
%lfs quota -u username /nobackup/username
Disk quotas for user username (uid xxxx):
   Filesystem  kbytes       quota   limit   grace   files   quota   limit   grace
/nobackup/username 1234  210000000 420000000    -      567   75000  100000       -
```

The NAS quota policy states that if you exceed the soft quota, an email will be sent to inform you of your current usage and how much of your grace period remains. It is expected that users will occasionally exceed their soft limit, as needed; however after 14 days, users who are still over their soft limit will have their batch queue access to Pleiades disabled.

If you anticipate having a long-term need for higher quota limits, please send a justification via email to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager for approval.

For more information, see also, Quota Policy on Disk Space and Files.

NOTE: If you reach the hard limit while your job is running, the job will die prematurely without providing useful messages in the PBS output/error files. A Lustre error with code -122 in the system log file indicates that you are over your quota.

In addition, when a Lustre filesystem is full, jobs writing to it will hang. A Lustre error with code -28 in the system log file indicates that the filesystem is full. The NAS Control Room staff normally will send out emails to the top users of a filesystem asking them to clean up their files.

Important: Backup Policy

As the names suggest, these filesystems are not backed up, so any files that are removed *cannot* be restored. Essential data should be stored on Lou1-3 or onto other more permanent storage.

Configurations

In the table below, /nobackupp[10-60] have been abbreviated as p[10-60].

Pleiades Lustre Configurations

Filesystem	p10	p20	p30	p40	p50	p60
# of MDSes	1	1	1	1	1	1
# of MDTs	1	1	1	1	1	1
size of MDTs	1.1T	1.0T	1.2T	0.6T	0.6T	0.6T
# of usable inodes on MDTs	~235x10^6	~115x10^6	~110x10^6	~57x10^6	~113x10^6	~123x10^6
# of OSSes	8	8	8	8	8	8
# of OSTs	120	60	120	60	60	60
size/OST	7.2T	7.2T	3.5T	3.5T	7.2T	7.2T
Total Space	862T	431T	422T	213T	431T	431T
Default Stripe Size	4M	4M	4M	4M	4M	4M
Default Stripe Count	1	1	1	1	1	1

NOTE: The default stripe count and stripe size were changed on January 13, 2011. For directories created prior to this change, if you did not explictly set the stripe count and/or stripe size, the default values (stripe count 4 and stripe size 1MB) were used. This means that files created prior to January 13, 2011 had those old default values. After this date, directories without an explicit setting of stripe count and/or stripe size adopted the new stripe count of 1 and stripe size of 4MB. However, the old files in that directory will retain their old default values. New files that you create in these directories will adopt the new default values.

Columbia Home Filesystems

DRAFT

This article is being reviewed for completeness and technical accuracy.

Columbia's home fileystem (/u/username) is NFS-mounted on the Columbia front-end (cfe2) and compute nodes (Columbia21-24).

Once a user is granted an account on Columbia, the home directory is set up automatically during his/her first login.

Quota and Policy

Disk space quota limits are enforced on the home filesystem. By default, the soft limit is 4GB and the hard limit is 5GB. There are no inode limits on the home filesystem.

To check your quota and usage on your home filesystem, do:

```
%quota -v
Disk quotas for user username (uid xxxx):
     Filesystem  blocks   quota   limit   grace   files   quota   limit   grace
   ch-rg1:/home6    4888  4000000 5000000            294       0       0
```

The quota policy for NAS states that if you exceed the soft quota, an email will be sent to inform you of your current usage and how much of your grace period remains. It is expected that a user will occasionally exceed their soft limit as needed; however after 14 days, users who are still over their soft limit will have their batch queue access to Pleiades disabled. If you believe that you have a long-term need for higher quota limits, you should send an email justification to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager, Bill Thigpen, for approval.

The quota policy for NAS can be found here.

Backup Policy

Files on the home filesystem are backed up daily.

Columbia CXFS Filesystems

Columbia CXFS filesystems (/nobackup[1-2][a-i]) are shared and accessible from cfe2 and Columbia21-24. This allows user jobs to be load-balanced across Columbia's systems without forcing users to move their data to a particular Columbia system.

Users will have a nobackup directory on one of these shared file systems. To find out where your nobackup directory is, log in to the front-end node and type the following shell command:

```
cfe2% ls -d /nobackup[1-2][a-i]/$USER
/nobackup1f/username/
```

In this example, the user is assigned to /nobackup1f.

Default Quota and Policy on /nobackup

Disk space and inodes quotas are enforced on the CXFS /nobackup[1-2][a-i] filesystems. The default soft and hard limits for inodes are 25,000 and 50,000, respectively. Those for disk space are 200GB and 400GB, respectively. To check your disk space and inodes usage and quotas on your CXFS filesystem, do the following:

```
cfe2% quota -v
Disk quotas for user username (uid xxxx):
     Filesystem  blocks   quota   limit   grace   files   quota   limit   grace
/dev/cxvm/nobackup1f
                1673856  210000000 420000000              10973   25000   50000
```

The NAS quota policy states that if you exceed the soft quota, an email will be sent to inform you of your current usage and how much of your grace period remains. It is expected that users will occasionally exceed their soft limit, as needed; however after 14 days, users who are still over their soft limit will have their batch queue access to Columbia disabled.

If you anticipate having a long-term need for higher quota limits, please send a justification via email to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager for approval.

For more information, see also, Quota Policy on Disk Space and Files.

Important: Backup Policy

As the names suggest, these filesystems are not backed up, so any files that are removed *cannot* be restored. Essential data should be stored on Lou1-3 or onto other more permanent storage.

Accessing CXFS from Lou

The Columbia CXFS filesystems are also mounted on Lou1-3. This allows you to copy files between the CXFS filesystems and your Lou home filesystem, using the *cp* or *cxfscp* commands on Lou.

Archive Systems

Mass Storage Systems: Lou1 and Lou2

The NAS environment contains three mass storage systems, Lou1 and Lou2, to provide long-term data storage for users of our high-end computing systems. These storage systems are SGI Altix computers running the Linux operating system. The disk space for the three systems combined is about 290 terabytes (TB), which is split into filesystems ranging from 9-30 TB in size.

Which Lou System I Should Use?

Each user should be able to log into any of the Lou systems, but will only have storage space on the home filesystem of one of them. Follow the steps below to determine which system you should store data on.

1. Log in to either Lou1 or Lou2. For example:

   ```
   your_localhost% ssh nas_username@lou1.nas.nasa.gov
   ```

2. Type the command "mylou" to find out your mass storage host. For example:

   ```
   lou1% mylou

   Your Mass Storage host is lou2

   Store files there in your home directory, /u/your_nas_username
   ```

   Be aware that Lou1 and Lou2 do *not* share their home filesystems.

3. Use the home filesystem on Lou*X* (where *X* = 1 or 2) determined by the step above for your long-term storage. For example:

   ```
   pfe1% scp foo lou2:
   ```

Quota Limits On Lou

For Lou*X* (where *X* = 1 or 2) that is assigned to you, there are no disk quota limits on your home filesystem. On the other hand, there *are* limits on the number of files (inode):

- 250,000 inode soft limit (14-day grace period)
- 300,000 inode hard limit

See Policy on Disk Files Quotas for Lou for more information.

Data (Un)Migration Between Disk and Tapes

In addition to the disk space, Lou1 and 2 have a combined 64 LTO-4 tape drives. Each of the LTO-4 tapes holds 800 GB of uncompressed data. The total storage capacity is approximately 10 PB.

Data stored on Lou's home filesystems (disk) is automatically migrated to tapes whenever necessary to make space for more data. Two copies of your data are written to tape media in silos located in separate buildings.

Data migration (from disk to tape) and unmigration (from tape to disk) are managed by the SGI Data Migration Facility (DMF) and Tape Management Facility (TMF).

If you need some data that is only available on tapes, make sure to unmigrate the data from tape to your home filesystem on Lou before transferring it to other systems.

For more tips on how to use Lou more effectively, see Storage Best Practices.

Network

TCP Performance Tuning for WAN Transfers

DRAFT

This article is being reviewed for completeness and technical accuracy.

The purpose of this document is to help you maximize your wide-area network bulk data transfer performance by tuning the TCP settings for your end hosts. These are some common configuration tasks for enabling high performance data transfers on your system.

Making changes to your system should only be done by a lead system administrator or someone who is authorized to make changes.

Linux

1. Edit */etc/sysctl.conf* and add the following lines:

   ```
   net.core.wmem_max = 4194304
   net.core.rmem_max = 4194304
   ```

2. Then have them loaded by running "sysctl -p".

Windows

We recommend using a tool like Dr. TCP

Set the "Tcp Receive Window" to at least 4000000, turn on "Window Scaling", "Selective Acks", and "Time Stamping".

Other options for tuning Windows XP TCP are the SG TCP Optimizer or using Windows Registry Editor to edit the registry, but this is only recommended for Windows users who are already familiar with registry parameters.

Mac

- **Do these steps for OS 10.4**

  These changes require root access.

In order to allow the Mac operating system to retain the parameters after a reboot, edit the following variables in */etc/sysctl.conf*:

```
# Set maximum TCP window sizes to 4 megabytes

net.inet.tcp.sendspace= 4194304
net.inet.tcp.recvspace= 4194304
# Set maximum Socket Buffer sizes to 4 megabytes

kern.ipc.maxsockbuf= 4194304
```

- **Do these steps for OS 10.5 and up**

  Use the **sysctl** command for the following variable:

  ```
  sysctl -w net.inet.tcp.win_scale_factor=8
  ```

If you follow these steps and are still getting less than your expected throughput, please contact the network group at support@nas.nasa.gov attn: Networks and we will work with you on tuning your system to optimize file transfers. You can also try the additional steps outlined in the following documents: Optional Advanced Tuning For Linux and Tips for File Transfers.

Optional Advanced Tuning for Linux

DRAFT

This article is being reviewed for completeness and technical accuracy.

This document describes additional TCP settings that can be tuned on high performance Linux systems. This is intended for 10 Gigabit hosts, but can also be applied to 1 Gigabit hosts. The following steps should be taken in addition to the steps outlined in TCP Performance Tuning for WAN transfers.

Configure the following */etc/sysctl.conf* settings for faster TCP:

Set maximum TCP window sizes to 12 megabytes

```
net.core.rmem_max = 11960320
net.core.wmem_max = 11960320
```

Set minimum, default, and maximum TCP buffer limits

```
net.ipv4.tcp_rmem = 4096 524288 11960320
net.ipv4.tcp_wmem = 4096 524288 11960320
```

Set maximum network input buffer queue length

```
net.core.netdev_max_backlog = 30000
```

Disable caching of TCP congestion state (Linux Kernel version 2.6 only). Fixes a bug in some Linux stacks.

```
net.ipv4.tcp_no_metrics_save = 1
```

Use the BIC TCP congestion control algorithm instead of the TCP Reno algorithm, Linux Kernel versions 2.6.8 to 2.6.18

```
net.ipv4.tcp_congestion_control = bic
```

Use the CUBIC TCP congestion control algorithm instead of the TCP Reno algorithm, Linux Kernel versions 2.6.18+

```
net.ipv4.tcp_congestion_control = cubic
```

Set the following to 1 (should default to 1 on most systems):

```
net.ipv4.tcp_window_scaling =1
net.ipv4.tcp_timestamps = 1
net.ipv4.tcp_sack = 1
```

A reboot will be needed for changes to */etc/sysctl.conf* to take effect, or you can attempt to reload sysctl settings (as root) with 'sysctl -p'.

For additional information visit this web site

If you have a 10Gig system or if you follow these steps and are still getting less than your expected throughput, please contact support@nas.nasa.gov attn: Networks and we will work with you on tuning your system to optimize file transfers.

NAS VPN Service

DRAFT

This article is being reviewed for completeness and technical accuracy.

For remote users wishing to connect to limited resources available on the local NAS network, a virtual private network service is available to any existing NAS user who has a Lou account and active SecurID fob. Additionally, NAS support staff may also make use of this service to provide some remote support to your government systems while on travel or at home.

*** ALL system traffic is routed through NAS while you are connected via VPN ***

This system is intended for government use and users are required to follow the appropriate use policy. All traffic is monitored. While connected, **ALL** your traffic will be routed through NAS, as if you were physically connected to NASLAN. When you are finished with your session, please remember to log out.

Users are subject to the NAS VPN Security Policy.

When and Why to use VPN:

The VPN service will make your system appear to be logically within the NAS external network, with some access to internal resources. Connection to the VPN server and installation of the VPN software are handled through Javascript and users do not have to pre-install prior to connection. VPN makes use of your standard web browser for encryption.

Through VPN, users can make use of any of the following services:

- Apple file sharing (AFP remote mount)
- Access to the NAS license server
- Access to the internal NAS webserver
- Access to some ARC websites not accessible from the outside
- SSH and SCP directly into NASLAN systems (NOT to Enclave systems)

How to Login to NAS VPN:

The login site is at: https://nas-vpn.nas.nasa.gov

Enter your login credentials (remember, it is your Lou account information and SecurID fob).

Click "Sign In", and you will be taken to the default VPN page upon successful authentication.

If it is the first time connecting from a new system, you will be prompted to install some software. Depending on your browser's security level, you may get a banner on top of the page warning you that the site is trying to install an ActiveX control. Click on that banner to install it.

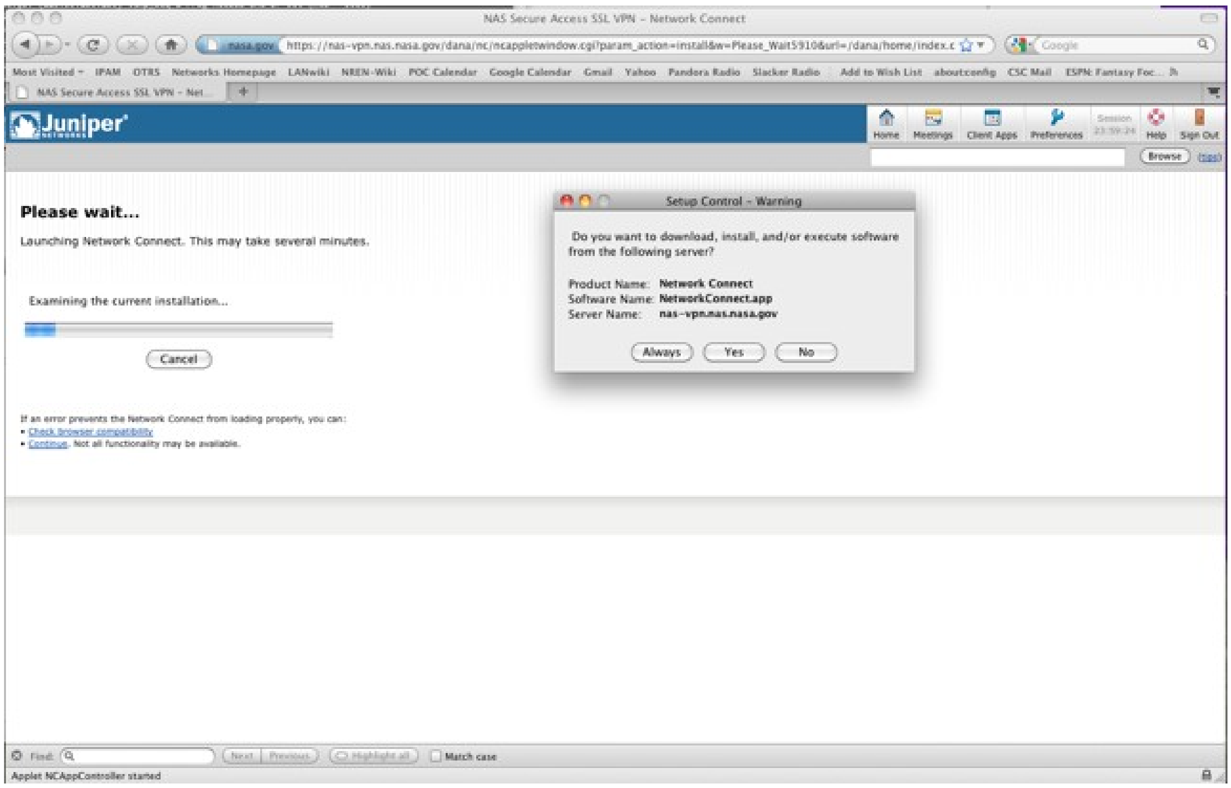
A pop-up window will show that the connection is negotiating. If it does not automatically start, click on "Start" under the Client Application Session for Network Connect. This may take up to a minute. Once done, the window will close and you will have been assigned a new IP address inside the VPN environment. A little icon will be placed in your desktop tray which you can use to get session information and disconnect.

You can verify that the VPN is working by connecting to:

http://myipaddress.com

Your address should be on the 198.9.30.x network.

NAS Supported systems already have the VPN Client installed on them. Just double click the "Network Connect" icon and enter your authentication credentials.



Please wait...

Launching Network Connect. This may take several minutes.

Examining the current installation...

If an error prevents the Network Connect from loading properly, you can:
- Check browser compatibility
- Continue. Not all functionality may be available.

Do you want to download, install, and/or execute software from the following server?

Product Name: Network Connect
Software Name: NetworkConnect.app
Server Name: nas-vpn.nas.nasa.gov

Applet NCAppController started

Alternatively, if you cannot connect through the above methods, you can manually download the VPN client software here:

- Download VPN client for Mac
- Download VPN client for Linux
- Download the tar file which contains the VPN client for Windows
- Download the tar file which contains the VPN client for Windows 64 bit systems

VPN FAQ's:

- DO I need to keep my web browser open to keep the VPN up?

  No, you do not need to keep your web browser open once you start up the Network Connect client.
- How do I disconnect from the VPN?

  You can either go back to http://nas-vpn.nas.nasa.gov and click the "Sign Out" button on the top right corner, or right-click on the icon in your system tray and select "Sign Out".
- Is there an auto-logout?

  Yes. You will be logged out after 30 minutes of inactivity. The max session length is 12 hrs before you need to re-authenticate and you will be given a reminder before being disconnected. This is so that people don't stay "camped" on the VPN network.
- What traffic is sent over the VPN?

  ALL traffic you send will be sent over the VPN. This includes any websites you visit, any chat programs, or any software that requires a network connection. Because of this, it is important that you disconnect from the VPN while not in use.
- What changes are made to my system?

  Several changes are made to your network including a new IP address, default route, search domain and other minor files which allow you to be "virtually" inside the NASLAN. This means you can refer to hosts just by their hostname and not their fully-qualified name - eg:

  ssh username@desktop
- What browsers are supported?

  As long as you meet the requirements listed above, you should be able to connect on Safari, Internet Explorer, and Firefox. We recommend you update to the latest stable version. The minimum browser requirements are:

  - ♦ 168-bit and greater encryption
  - ♦ SSLv3 and TLSv1
  - ♦ JRE / Java enabled
  - ♦ Pop-up Windows

- How will connecting to the VPN impact my home network?

Some of your home services may stop working while connected to the VPN. This includes services like Internet printing, file sharing, and audio streaming. This is to ensure security of NAS while you are connected to the VPN.

NAS Remote Network Diagnostic Tools

A NAS network service that enables remote users to test end-to-end connectivity to the NAS HECC enclave is now available. Users can access the Network Diagnostic Tool service and initiate tests at: http://npad.nas.nasa.gov.

The diagnostic tests can only run on a Java-enabled web browser. If you have trouble accessing the website, please contact the NAS Control room (see below) and we will assist you.

Features

- Tools are accessible from any standard web browser
- Command-line tools are also available for Linux servers
- Users receive a diagnostic report on the test results; a copy of the report is sent to the NAS Networks team for analysis. If any problems are identified, the team will contact you to help resolve the issue.

The services available on this website run from inside the NAS network and are connected at 10 Gigabit Ethernet rates.

The Network Diagnostic Tester (NDT) - Performs a quick test that reports the maximum throughput to your remote system from NAS. It will also identify any issues with possible bad cabling, negotiation issues, packet loss, or general network congestion.

Network Path Application Diagnosis (NPAD) tool - Performs a more elaborate connection test and determines problems with TCP parameters, buffer sizes, and/or router congestion, and notifies you of recommended settings for maximum performance.

Traceroute tool - Allows users to perform reverse traceroutes from NAS display the path and measure transit delays of packets to a remote host.

Ping tool - Allows users to perform reverse pings from NAS to the remote host, to test the reachability to your host and measure round-trip time for messages to reach the remote host.

If you want a command line interface, you can also download client software and link to various other public services from this website.

If you have any questions about these new services, please contact the NAS Control Room staff 24x7 at (800) 331-8737, (650) 604-4444, support@nas.nasa.gov.

Increasing File Transfer Rates

One challenge users face is moving large amounts of data efficiently to/from NAS across the network. Often, minor system, software, or network configuration changes can increase network performance an order of magnitude or more. This article describes some methods for increasing data transfer performance.

If you are experiencing slow transfer rates, try these quick tips:

- Transfer using the bridge nodes (bridge1, bridge2) instead of the Pleiades front-end systems (PFEs). The bridge nodes have much more memory, along with 10-Gigabit Ethernet interfaces to accommodate many large transfers. The PFEs often become oversubscribed and cause slowness.
- If using the scp command, make sure you are using OpenSSH version 5 or later. Older versions of SSH have a hard limit on transfer rates and are not designed for WAN transfers. You can check your version of SSH by running the command ssh -V.
- For large files that are a gigabyte or larger, we recommend using BBFTP. This application allows for transferring simultaneous streams of data and doesn't have the overhead of encrypting all the data (authentication is still encrypted).

Online Network Testing Tools

The NAS PerfSONAR Service provides a custom website that that allows you to quickly self-diagnose your remote network connection issues, and reports the maximum bandwidth between sites, as well as any problems in the network path. Command-line tools are available if your system does not have a web browser.

Test results are also sent to our network experts, who will analyze traffic flows, identify problems, and work to resolve any bottlenecks that limit your network performance, whether the problem is at NAS or a remote site.

One-on-One Help

If you still require assistance in increasing your file transfer rates, please contact the NAS Control Room at support@nas.nasa.gov, and a network expert will work with you or your local administrator one-on-one to identify methods for increasing your rates.

To learn about other network-related support areas. see also, End-to-End Networking Services.

Auxiliary Systems

Bridge nodes

DRAFT

This article is being reviewed for completeness and technical accuracy.

Currently, the Pleiades Lustre filesystems are not mounted on Lou. File transfers between Pleiades and Lou are normally done with remote file transfer commands such as scp, bbftp and bbscp.

Using the Pleiades bridge nodes, one can actually transfer files between Pleiades' home or Lustre filesystems and Lou's home filesystem through Columbia's CXFS filesystems.

In the example below, on bridge1, the file *foo* is copied from a user's Pleiades Lustre /nobackup filesystem to his Columbia CXFS filesystem under /nobackup2a. Then, on Lou, the same file is copied from /nobackup2a to user's Lou home filesystem.

```
bridge1% cp /nobackup/username/foo /nobackup2a/username
```

```
lou1% cp /nobackup2a/username/foo /u/username
```

File Transfers

File Transfer: Overview

Here's a general overview of the various file transfer scenarios within the NAS environment, with pointers to related articles.

File Transfers Between Pleiades, Columbia, and Lou

File transferring between NAS systems in the secure enclave (Pleiades, Columbia, and Lou) uses host-based authentication (transparent to users) and is usually straightforward. The following articles provide basic information to help you get started.

- Local File Transfer Commands - cp, cxfscp
- Remote File Transfer Commands - scp, bbftp/bbscp
- File Transfer Between Pleiades and Columbia or Lou
- Transferring Files from the Pleiades Compute Nodes to Lou
- Checking File Integrity

File Transfers between a NAS HECC Host and Your Localhost

Transferring files between a NAS host (such as Pleiades,Columbia, or Lou) and a remote host, such as your local desktop, is more complex. There are multiple factors that you should be aware of:

Which commands to use

Remote File Transfer Commands such as *scp* and *bbftp* and *bbscp* are supported on most NAS high-end computing systems. Depending on the way the transfers are performed, you may need either one or both of the client and server software of scp and/or bbftp or the bbscp script installed on your localhost.

- **Transfer Rate**

  File transfer rate with scp, especially using scp from versions of Open that SSH are older than 4.7, can be as slow as 2 MB/sec. For transferring large files over a long distance, consider the following:

  - ♦ upgrade to the the latest version of OpenSSH
  - ♦ apply the HPN-SSH patch to your OpenSSH

    - ♦ enable compression by adding -C to the scp command-line if the data will compress well
    - ♦ use bbftp/bbscp

- **Security Issues**
    - ♦ With scp, users' authentication information (such as password or passcode) and data are encrypted.
    - ♦ With bbftp and bbscp, only the authentication information is encrypted, while data is not.
    - ♦ You can use GPG to encrypt your data prior to the transfer.

Where transfer commands are initiated

- **Outbound file transfers**

  When the file transfer command is initiated on a NAS host such as Pleiades, Columbia, or Lou, the transfer does not need to go through SFE[1,2] or Secure Unattended Proxy. This is the easiest way to transfer files from and/or to your site if your localhost is configured to allow the transfer.

  To learn more, see also Outbound File Transfer Examples.

- **Inbound file transfers**

  When the file transfer command is initiated on a remote host such as your local desktop, the transfer must go through either SFE[1,2] or Secure Unattended Proxy.

    - ♦ *Going through SFE[1,2]*

      Going through SFE[1,2] requires authentication via your RSA SecurID fob at the time of operation; you will be prompted for your passcode when you issue the file transfer commands, such as *scp*, *bbftp*, or *bbscp*.

      Transfers can be done with one of the following two approaches:

        1. Two steps: Initiate scp from your localhost to SFE[1,2], and then initiate another scp from SFE[1,2] to Columbia, Pleiades or Lou.

           WARNING: Do not store files on the SFEs since space is very limited. Any file transfers though the SFE really should use the SSH pass-through option described next.
        2. One step: Initiate scp, bbftp/bbscp from your localhost to PleiadesmColumbia, or Lou if SSH Passthrough has been set up.

      To learn more, see also Inbound File Transfers through SFEs Examples.

♦ *Going through SUP*

Going through the Secure Unattended Proxy does *not* require SecurID fob authentication at the time of operation. Instead, special "SUP keys" using SecurID authentication must be obtained ahead of time. The "SUP keys" are good for one week and are used automatically to authenticate users for file transfers using scp, bbftp or bbscp issued on a command-line or in a job script.

WARNING: Although users have accounts on the SUP servers, no login session is allowed.

File transfers going through SUP offers multiple benefits over going through the SFEs:

◊ SUP allows the transfer to be unattended; that is, you do not have to type in your password, passphrase, or passcode when the file transfer command is issued. So, file transfers can be done within a script that can be scheduled to run ahead of time. On the other hand, file transfers through the SFEs can not be done in a script.
◊ File transfers through SUP are done in one step, and setting up SSH passthrough is not needed since the SFEs are not involved.
◊ SUP automatically sets some options, such as the port range allowed for bbftp transfers, so that you don't have to set them explicitly. Thus, the syntax for bbftp over SUP is greatly simplified compared to bbftp without SUP.

NOTE: Some sites only allow specific outbound ports; this may cause bbftp to break.
◊

See the article Using the Secure Unattended Proxy (SUP) and the examples there for more information.

• **File staging**

When there are issues (such as a firewall) that hinder the inbound and/or outbound transfers, file staging through DMZFS[1,2] is another option. You can deposit and retrieve files on DMZFS[1,2] by issuing the *scp* or *bbftp* command on either a NAS host or your localhost.

WARNING: The total storage space on DMZs is 2.5TB, shared among all users; files older than 24 hours are removed.

DMZFS[1,2] do not use SecurID fob for authentication. Instead, password or public key authentication is used for file transfers via DMZFS[1,2].

Unattended file transfers can also be done through DMZFS[1,2] if public key authentication has been set up on DMZFS[1,2].

Note, however, that for this purpose, the SUP is preferred as SUP transfers are direct to the end target so do not have the storage restrictions and two step performance limitations of DMZFS when using bbftp/bbscp.

Read File Staging through DMZ File Servers for more information.

NAS Username and Your Local Username

If your NAS username and local username are different, you may have to add the appropriate username in the scp, bbftp or bbscp command-line.

- If you issue the command on your local host, then the username is your NAS username.
- If you issue the command on a NAS host, then the username is your local username.

In the examples shown in the articles Outbound File Transfer Examples and Inbound File Transfers through SFEs Examples, you will find the correct syntax for adding the appropriate username in the file transfer commands.

For inbound file transfers, if you have correctly included your NAS username in the *~/.ssh/config* file of your localhost, you do not have to include the NAS username in the *scp*, *bbftp* or *bbscp* command. A template for the *~/.ssh/config* is available for download.

Check File Integrity Before and After the Transfer

It's a good practice to ensure the integrity of the data before and after the transfer. For more information, see Checking File Integrity.

Tuning your Local System to Improve File Transfer Performance

Some file transfer commands provide options that can be used to improve your transfer rates. For example, enabling compression during file transfers may help in some cases: with bbftp, you can use multiple streams instead of a single stream for better performance. Read Tips for File Transfers for more information.

On the other hand, file transfer performance is also dependent on some system-wide settings. If necessary, ask your local system administrator to look into issues discussed in the following articles:

- TCP Performance Tuning for WAN Transfers
- Optional Advanced Tuning for Linux
- Pittsburgh Supercomputing Center's Enabling High Performance Data Transfers - a properly tuned TCP/IP stack

Local File Transfer Commands

DRAFT

This article is being reviewed for completeness and technical accuracy.

The following file transfer commands can be used when both the source and destination locations are accessible on the same host where the command is issued. Basic information about each command is provided below.

- **cp**:

  cp is a UNIX command for copying files between two locations (for example, two different directories of the same filesystem or two different filesystems such as NFS, CXFS or Lustre).

  - *Where is it installed at NAS?*

    cp is available on all NAS systems except SFE[1,2], DMZFS[1,2], Bouncer and Bruiser.

  - *Examples:*

    ```
    pfe1% cp $HOME/foo $HOME/newdir/foo2
    pfe1% cp $HOME/foo /nobackup/username
    ```

- **cxfscp**:

  cxfscp is a program from SGI for quickly copying large files to and from a CXFS filesystem (for shared-memory systems such as Columbia). It can be significantly faster than cp on CXFS filesystems since it uses multiple threads and large direct I/Os to fully utilize the bandwidth to the storage hardware.

  For files less than 64 kilobytes in size, which will not benefit from large direct I/Os, cxfscp will use a separate thread for copying these files using buffered I/O similar to cp.

  - *Where is it installed at NAS?*

    cxfscp is installed on cfe2, c21-24, Lou[1-2], and the Pleiades bridge nodes (bridge[1-2]). It is not available on the Pleiades front-end nodes (pfe[1-12]).

  - *When to use it?*

    The Columbia CXFS filesystems (/nobackup[1-2][a-i]) are mounted on all Columbia hosts (cfe2, c21-24), Lou[1-3], and the Pleiades bridge nodes (bridge[1,2]). The command cxfscp can be issued on any of these hosts to copy large files to and from Columbia's /nobackup[1-2][a-i]. This is an easy way to transfer files among Columbia, Pleiades and Lou without the need for scp, bbftp or bbscp.

- *Examples:*

```
cfe2% cxfscp /nobackup2a/username/foo /nobackup2a/username/new_dir
lou2% cxfscp /nobackup2a/username/foo $HOME
bridge2% cxfscp $HOME/foo /nobackup2a/username
bridge2% cxfscp /nobackupp20/username/foo /nobackup2a/username
```

- *Performance:*

  Some benchmarks done by NAS staff show that cxfscp is typically 4 - 7 times faster than cp for large files (2+ GB) and can achieve upto 400 MBytes/sec.

For more information, read **man cxfscp**.

Remote File Transfer Commands

DRAFT

This article is being reviewed for completeness and technical accuracy.

The following file transfer commands can be used when the source and destination are located at different hosts. They can be used to transfer files either between NAS HECC hosts or between a NAS host and a remote host such as your local desktop system.

- **scp (with/without HPN-SSH patch)**

  Secure Copy Protocol (SCP), based on Secure Shell Protocol (SSH), is a means of securely transferring files between a local and a remote host. Both the authentication information (such as password or passcode) and user's data are encrypted.

  Normal scp (without the HPN-SSH patch)

  The most widely used scp is from OpenSSH.

  - ♦ *Where is it installed at NAS?*

    A copy of scp from OpenSSH without the patch is available on the Pleiades front-end and bridge nodes (pfe[1-12], bridge[1-2]), all Columbia nodes, Lou[1-2], SFE[1,2], Bouncer, and Bruiser.

    The copy on SUP contains the HPN-SSH patch.

    scp is not available on DMZFS[1,2]. Use scp on Columbia, Pleiades, Lou or your localhost to push files into DMZFS[1,2] or pull files out of DMZFS[1,2].

  - ♦ *Do you need it installed on your localhost?*

    If you already have a version of SSH installed on your localhost, most likely, scp is already there.

  - ♦ *When to use it?*

    scp is typically used for transferring small files (<< 5GB) within NAS or offsite (<< 1 GB) that takes reasonable amount of time to complete.

  - ♦ *Examples:*

    For outbound transfer:

lou1% scp local_username@your_localhost.domain:foo ./foo2

For inbound two-step transfer:

```
your_localhost% scp foo nas_username@sfe1.nas.nasa.gov:foo2
sfe1% scp foo2 lou1:
```

For inbound one-step transfer if SSH-passthrough has been set up correctly:

your_localhost% scp foo nas_username@lou1.nas.nasa.gov:foo2

To transfer files through DMZFS[1,2], initiate the scp command from either a NAS HECC host or your localhost, not DMZFS[1,2]. For example,

```
your_localhost% scp foo nas_username@dmzfs1.nas.nasa.gov:foo2
pfe1% scp dmzfs1:foo2 .
```

Omit *local_username@* and *nas_username@* in the examples above if your local username and NAS username are identical.

♦ *Performance:*

  ◊ Within NAS HECC Enclave, depending on source and destination hosts and other factors, the performance range will be 40 - 100 Mbytes/sec.

  ◊ Over WAN (such as between NAS and a remote site), the best you get with scp from OpenSSH versions older than 4.7 (with the internal channel buffer set to 128 KB) is ~ 2 MBytes/sec. Starting with OpenSSH version 4.7, a larger channel buffer is introduced to improve file transfer performance. Users are recommended to upgrade to version 4.7 or later.

    In case where OpenSSH 4.7 or a later version does not yield satisfactory performance, consider applying the HPN-SSH patch to your OpenSSH.

    If the data you are transferring will compress well, consider enabling compression by adding -C to your scp command-line.

HPN-SSH enabled scp

HPN-SSH is a patch for OpenSSH designed to eliminate a network throughput bottleneck that typically occurs in an SSH session over long distance and high bandwidth network (i.e.,when the bandwidth-delay product is high). This is accomplished by allowing internal flow control buffers to be defined and grow at runtime, rather than statically defining them as OpenSSH does. The resulting performance increase can range from 10x to more than 50x, depending on the cipher used and host tuning.

HPN-enabled SSH is fully interoperable with other SSH servers and clients. HPN clients will be able to download faster from non-HPN servers, and HPN servers will be able to receive uploads faster from non-HPN clients. However, the host receiving the data must have a properly tuned TCP/IP stack.

Ask your local network staff for help to see if an HPN-SSH patch is needed for certain network connection.

♦ *Where is it installed at NAS?*

  ◊ On cfe2, the client version of OpenSSH 4.7p1 with HPN12v20 patch is available.
  ◊ On Lou[1-2], the client version of OpenSSH 5.0p1 with HPN13v1 patch is available.
  ◊ On SUP, both the client and server of OpenSSH 5.1p1 have been patched with HPN13v5.

On cfe2 and Lou[1-2], the HPN-patched SSH programs are purposely named as *hpn-ssh, hpn-scp,* and *hpn-sftp* to distinguish them from the default non-HPN versions (*ssh, scp* and *sftp*) of OpenSSH.

♦ *Do you need it installed on your localhost?*

To get good performance, an HPN-SSH server must be installed on your local system if data is to be received on your local system.

Typical installation procedure:

1. Download OpenSSH source (*openssh-x.xpx.tar.gz*) from http://www.openssh.com
2. Download corresponding HPN SSH patch (*openssh-x.xpx-hpnxxvx.diff.gz*) from http://www.psc.edu/networking/projects/hpn-ssh
3. Uncompress and untar above distributions
4. move the file *openssh-x.xpx-hpnxxvx.diff* to the directory *openssh-x.xpx*
5. cd openssh-x.xpx (for example, openssh-5.0p1)
6. patch < openssh-5.0p1-hpn.diff
7. configure [OPTIONS]
8. make [OPTIONS]
9. Validate:
```
%ssh -v
OpenSSH_5.0p1-hpn13v3
```

♦ *Examples:*

```
lou[1-2]% hpn-scp -c arcfour -o TCPRcvBufPoll=yes source destination
your_localhost% scp -c arcfour -o TCPRcvBufPoll=yes source destination
```

Note:

- ◊ arcfour (RC4) is a more CPU-efficient 128-bit cipher. One can also choose NONE for cipher so that there is no encryption for data.
- ◊ Enabling *TCPRcvBufPoll* allows for the TCP receive buffer to be polled through the duration of the connection.

♦ *Performance:*

With an HPN-SSH enabled scp, one can expect good performance for transferring large files to remote sites over long distance with high latency connections. Improvement over non-patched scp older than 4.7 (2 Mbytes/sec) may be 10x to 50x.

• **bbFTP**

bbFTP is a high performance remote file transfer protocol which supports parallel TCP streams for data transfers. Basically, it splits a single file in several pieces and sends them through parallel streams. The whole file is then rebuilt on the remote site. bbFTP also allows dynamically adjustable TCP/IP window sizes instead of a statically defined window size used by normal scp. In addition, it provides a secure control channel over SSH and allows data to be transferred in cleartext to reduce overhead in unnecessary encryption. These characteristics allow bbftp to achieve bandwidths that are greater than normal scp.

bbFTP is recommended in place of *scp* for the data transfer of large files over long distances.

♦ *Where is it installed at NAS?*

Both the bbFTP server (*bbftpd*) and client (*bbftp*) are installed on all Columbia hosts, Lou[1-2], Pleiades front-end and bridge nodes (pfe[1-12], bridge[1-2]) and SUP.

For DMZFS[1,2], only the bbFTP server (*bbftpd*) is installed. Issue the *bbftp* command from Columbia, Pleiades, Lou or your localhost (if bbFTP client has been installed) to push files into DMZFS[1,2] or pull files out of DMZFS[1,2].

♦ *Do you need it installed on your localhost?*

If you want to initiate *bbftp* from your localhost, you have to download and install the client version of bbFTP on your localhost. If you want to initiate *bbftp* from a NAS HECC system and transfer files from/to your localhost, download and install the server version of bbFTP on your localhost.

♦ *When to use it?*

Consider using bbFTP when transferring large files (> 1 GB) within NAS or offsite. Be sure to use multiple streams to get better transfer rate.

- *Example:*

  bbftp is like a non-interactive ftp and the syntax can be complicated.

  ```
  your_localhost% bbftp -u nas_username -e 'setnbstream 8; get filename'
                  -E 'bbftpd -s -m 8' lou1.nas.nasa.gov
  ```

  ```
  For formatting issue, the above command was broken into
  two lines. In reality, it should be just one line.
  ```

- *Performance:*

  ◊ bbFTP typically transfers data 10 - 20 times faster than normal scp.

  ◊ Within NAS HECC Enclave, performance should be 100+ MB/sec.

  ◊ Over WAN, the performance can be upto 50 MBytes/sec. File transfers between NAS and certain NASA sites may reach 100 Mbytes/sec.

  If you are not getting good performance, check with your network administrator to see if performance tuning is needed on your system.

The article bbFTP provides more instructions on installing and using bbFTP.

- **bbSCP**

bbSCP is written in Perl by Greg Matthews at NAS. It is a bbftp wrapper which provides an scp-like command-line interface. It assembles the proper command-line for bbftp and then executes bbftp to perform the transfers. bbSCP is designed and tested for bbftp version 3.2.0.

bbSCP only encrypts usernames and passwords, it does NOT encrypt the data being transferred.

- *Where is it installed at NAS?*

  bbSCP is installed on all Columbia hosts, Lou[1-2], Pleiades front-end and bridge nodes (pfe[1-12], bridge[1-2]) under /usr/local/bin.
- *Do you need it installed on your localhost?*

  If you want to initiate bbscp from your localhost, you need to:

  ◊ download and install bbftp-client-3.2.0 on your localhost
  ◊ download bbscp version 1.0.6 (also attached at the end of this article) and install it on your localhost
- *When to use it?*

Use the bbSCP script when you want the bbftp functionality and performance but with scp-like syntax. It can be used for transferring files within NAS HECC Enclave or between NAS and a remote site.

- *Example:*

```
your_localhost% bbscp foo nas_username@lou1.nas.nasa.gov:
```

- *Performance:*

  Performance of bbSCP is the same as bbFTP.

The article bbscp provides more information (man page, performance turing, test and verification).

Outbound File Transfer Examples

DRAFT

This article is being reviewed for completeness and technical accuracy.

When the file transfer command (such as scp, bbftp or bbscp) is initiated on a NAS HECC host such as Columbia, Pleiades or Lou, the transfer does not need to go through SFE[1,2] or SUP. This is the fastest way to transfer files from/to your site if your localhost is configured to allow the transfer.

To simplify the instructions, the approaches will be described in terms of transfers to/from one of the Pleiades front-end node, pfe1, but they also apply to any of the other systems that are in the enclave (such as other Pleiades front-end or bridge nodes, Columbia or Lou). For each method described, two commands are provided. The first one is used when the user have identical username between his/her localhost and the NAS HECC systems. The second one is used when the usernames are different.

Logging into pfe1 and

- *Using scp for the outbound transfer:*

  To push files out of pfe1,

  ```
  pfe1% scp foo your_localhost:
  pfe1% scp foo local_username@your_localhost:
  ```

  To pull files into pfe1,

  ```
  pfe1% scp your_localhost:foo .
  pfe1% scp local_username@your_localhost:foo .
  ```

- *Using bbftp for outbound transfer:*

  If you find that using scp gives poor performance rates, we recommend using the application bbftp. This will require that the bbFTP server (bbftpd) is installed on your localhost.

  To push files out of pfe1,

  ```
  pfe1% bbftp -s -e 'setnbstream 8; put foo' your_localhost
  pfe1% bbftp -s -u local_username -e 'setnbstream 8; put foo' your_localhost
  ```

  To pull files into pfe1,

  ```
  pfe1% bbftp -s -e 'setnbstream 8; get foo' your_localhost
  ```

```
pfe1% bbftp -s -u local_username -e 'setnbstream 8; get foo' your_localhost
```

See bbftp for more instructions.

- *Using bbscp for outbound transfer:*

  bbSCP is a wrapper for bbFTP which provides scp-like syntax. Using this method for outbound transfer requires that the bbFTP server (bbftpd) is installed on your localhost.

  To push files out of pfe1,

  ```
  pfe1% bbscp foo your_localhost:
  pfe1% bbscp foo local_username@your_localhost:
  ```

  To pull files into pfe1,

  ```
  pfe1% bbscp your_localhost:foo .
  pfe1% bbscp local_username@your_localhost:foo .
  ```

See bbscp for more instructions.

Inbound File Transfer through SFEs Examples

DRAFT

This article is being reviewed for completeness and technical accuracy.

Inbound file transfers through SFEs require SecurID fob authentication, and the transfer can be done in two steps or one step depending on whether you have set up SSH passthrough.

To simplify the instructions, the approaches will be described in terms of transfers to/from one of the Pleiades front-end node, pfe1, but they also apply to any of the other systems that are in the enclave (such as other Pleiades front-end or bridge nodes, Columbia or Lou). For each method described, two commands are provided. The first one is used when (1) the user have identical username between his/her localhost and the NAS HECC systems, or (2) the usernames are different but the user has set up his/her local *~/.ssh/config* file to include the NAS username. To learn how to set this up, download the ~/.ssh/config template. The second one is used when the usernames are different and the user does not include the NAS username in his/her local *~/.ssh/config* file.

- Two-step file transfers

  If you have not set up SSH passthrough, this will be your only option for inbound file transfers. It requires you to use scp twice: once on your localhost to transfer files to/from one of the SFEs (for example, sfe1), and the second one on the SFE or the host inside the HECC Enclave to transfer files between SFEs and the HECC host such as pfe1.

  To push files out of your localhost,

  ```
  step 1:
  your_localhost% scp foo sfe1.nas.nasa.gov:
  your_localhost% scp foo nas_username@sfe1.nas.nasa.gov:

  step 2:
  sfe1% scp foo pfe1:
  or
  pfe1% scp sfe1:foo .
  ```

  To pull files into your localhost,

  ```
  step 1:
  sfe1% scp pfe1:foo .
  or
  pfe1% scp foo sfe1:

  step 2:
  ```

```
your_localhost% scp sfe1.nas.nasa.gov:foo .
your_localhost% scp nas_username@sfe1.nas.nasa.gov:foo .
```

- One-step file transfers

  If you have set up SSH passthrough correctly, you can use either scp, bbftp or bbscp to transfer files between your localhost and a NAS HECC host.

  - *Using scp,*

    To push files out of your localhost,

    ```
    your_localhost% scp foo pfe1.nas.nasa.gov:
    your_localhost% scp foo nas_username@pfe1.nas.nasa.gov:
    ```

    To pull files into your localhost,

    ```
    your_localhost% scp pfe1.nas.nasa.gov:foo .
    your_localhost% scp nas_username@pfe1.nas.nasa.gov:foo .
    ```

  - *Using bbftp,*

    This requires that you have a bbftp client installed on your localhost.

    To push files out of your localhost,

    ```
    your_localhost% bbftp -s -e 'setnbstream 8; put foo' pfe1.nas.nasa.gov
    your_localhost% bbftp -s -u nas_username
                    -e 'setnbstream 8; put foo' pfe1.nas.nasa.gov
    ```

    ```
    For formatting issue, the second command above was broken
    into two lines. In reality, it should be in one line.
    ```

    To pull files into your localhost,

    ```
    your_localhost% bbftp -s -e 'setnbstream 8; get foo' pfe1.nas.nasa.gov
    your_localhost% bbftp -s -u nas_username
                    -e 'setnbstream 8; get foo' pfe1.nas.nasa.gov
    ```

    ```
    For formatting issue, the second command above was broken
    into two lines. In reality, it should be in one line.
    ```

    See bbftp for more instructions.
  - *Using bbscp,*

    This requires that you have the bbftp client version 3.2.0 and the NAS bbscp script installed on your localhost.

To push files out of your localhost,

```
your_localhost% bbscp foo pfe1.nas.nasa.gov:
your_localhost% bbscp foo nas_username@pfe1.nas.nasa.gov:
```

To pull files into your localhost,

```
your_localhost% bbscp pfe1.nas.nasa.gov:foo .
your_localhost% bbscp nas_username@pfe1.nas.nasa.gov:foo .
```

See bbscp for more instructions.

Using the Secure Unattended Proxy (SUP)

The Secure Unattended Proxy (SUP) allows users to perform remote operations on specific hosts within the HEC enclave (currently the Columbia front-ends, Pleiades front-ends/bridge nodes, Lou[1-2], and Susan)

without

the use of SecurID at the time of the operation. Users must obtain special "SUP keys" using SecurID authentication, after which those keys can be used to perform operations from unattended jobs and/or scripts.

SUP keys are currently allowed to call *scp, sftp, bbftp, qstat, rsync,* and *test*. In the future, other operations may be available via the SUP. Each SUP key is valid for a period of **one week** from the time it is generated. Users may have multiple SUP keys at the same time, which will expire asynchronously.

SUP Usage Summary

The steps below demonstrate how to quickly get up and running with the SUP using an scp transfer to pfe1 as an example. Consult the link in each step for full details (or simply read this page to completion).

1. Download and install client (one time)

   ```
   your_localhost% wget -O sup http://hecc.nas.nasa.gov/kb/file/9
   your_localhost% chmod 700 sup
   your_localhost% mv sup ~/bin
   ```

2. Authorize host for SUP operations (one time per host)

   ```
   your_localhost% ssh pfe1
   pfe1% touch ~/.meshrc
   ```

3. Authorize directories for writes (one or more times per host)

   ```
   your_localhost% ssh pfe1
   pfe1% echo /tmp >>~/.meshrc
   ```

4. Execute command (each time)

   ```
   your_localhost% sup scp foobar pfe1:/tmp/c_foobar
   ```

5. Examine expected output (as needed)

6. Troubleshoot problems (as needed)

SUP Client

The SUP client performs all the steps needed to execute commands through the SUP as if the SUP itself did not exist. Commands that are allowed to pass through the SUP can be executed as if the remote host were directly connected by simply prepending the client command "sup". Besides executing remote commands, the client also includes an operating system-independent virtual file system that allows files across all SUP-connected resources to be accessed using standard filesystem commands.

- **Requirements**

  The client requires Perl version 5.6.1 or above to execute and has been tested successfully on Linux, OS X, and Windows under Cygwin and coLinux. Only SSH is required to use the SUP, however, so if these requirements cannot be met, it is possible to use the SUP without the client.

  Note for Windows users: even if the client is not used, scp and sftp require functionality only found in the OpenSSH versions of these commands, so Cygwin or coLinux will still be needed.
- **Installation**

  1. Download the client and save to a file called "sup"

  2. Make the client executable using "chmod 700 sup"

  3. Move the client to a location in your $PATH

- **SSH Configuration**

  If your local username differs from your NAS username, it is recommended that you add the following to your *~/.ssh/config* file, where "nas_username" should be replaced with your NAS username:

  ```
  Host sup.nas.nasa.gov sup-key.nas.nasa.gov
       User nas_username
  ```

  NOTE: If you are using a config file based on the NAS config template, you do not have to do this step.

  Alternatively, the client's -u option can be used as described in the next section. If your local username is the same as your NAS username, no additional configuration or command-line options are required.

- **SUP Command-line Options**

  ♦ -b

Disable user interaction for use within scripts. Note that the client will fail if any interaction is required - normally only needed when your SUP key has expired or is otherwise unavailable.

- -k

  By default, the client leaves any SSH agents started on your behalf running for future invocations after the client exits. This option forces spawned agents to be killed before exiting. Note that "-b" automatically implies "-k".

- -u user

  Specify NAS username. Note that this option is required if your local username differs from your NAS username and you have not modified your SSH configuration appropriately.

- -v

  Enable verbose output for debugging purposes.

SUP Authorizations

The basic set of operations that may be performed using the SUP is specified by the administrator. To protect accounts from malicious use of SUP keys, users must grant execute and write permissions to SUP operations on each target system.

1. **Execute Authorization**

   By default, even SUP operations permitted by site policy are not allowed to execute on a given host. To enable SUP operations to a given host (currently, the Columbia front-ends, Pleiades front-ends/bridge nodes, Lou[1-2], or Susan), the file *~/.meshrc* must exist on that host, which can be created by invoking the following:

   ```
   touch ~/.meshrc
   ```

   Note that the Pleiades front-ends/bridge nodes share their home filesystems, so this must only be done on one of these nodes. Similarly, the Columbia front-ends share their home filesystems and the *~/.meshrc* file only needs to be created on one of the Columbia front-end nodes. Other systems must be authorized separately. Once this file exists on a host, all operations permitted by site policy are allowed to execute on that host.

2. Write Authorization

By default, SUP operations are not allowed to write to the file system on a given host. To enable writes to a given directory on a given host, that directory must be added (on a separate line) to the *~/.meshrc* file on that host. For example, the following lines in *~/.meshrc* indicate that writes should be permitted to /nobackupp40 and /tmp.

```
/nobackupp40
/tmp
```

Each directory is the root of allowed writes, so this configuration would allow writes to all files and directories rooted at /nobackupp40 and /tmp (for example, /nobackupp40/some/dir, /tmp/some/file).

Note that the root directory cannot be authorized. Also note that dot files (i.e. *~/.**) in your home directory are never writable regardless of the contents of *~/.meshrc*.

Executing Commands Through SUP

Usage example of each command that may be executed through the SUP are given below. Note that SUP commands must be authorized for execution on each target host, and that transfers to a given host must be authorized for writes. Before a given operation is performed, the client may ask for certain information, including the existing or new passphrase for *~/.ssh/id_rsa*, the password + passcode for sup.nas.nasa.gov, and/or the password + passcode for sup-key.nas.nasa.gov.

File Transfer Commands

bbftp (man page)

```
your_localhost% sup bbftp -e "put foobar /tmp/c_foobar"
pfe1.nas.nasa.gov
```

Note that you must use the fully qualified domain name of the target host (in this case, pfe1.nas.nasa.gov) if you are not within the NAS domain.

bbscp (man page)

```
your_localhost% sup bbscp foobar pfe1.nas.nasa.gov:/tmp/c_foobar
```

Note that bbscp is just a client-side wrapper for bbftp, therefore, as with bbftp, you must use the fully qualified domain name of the target host (in this case, pfe1.nas.nasa.gov) if you are not within the NAS domain.

rsync (man page)

```
your_localhost% sup rsync foobar pfe1:/tmp/c_foobar
```

If you intend to transfer files to your home directory, note that even if your home directory has been authorized for writes, **rsync transfers to your home directory will fail unless the "-T" or "--temp-dir" option is specified**. This is because rsync uses temporary files starting with "." during transfers, which cannot be written in your home directory. You can avoid this problem by specifying an alternate temporary directory that is authorized for writes. For example, the following example uses /tmp as the temporary directory when files are transferred to the home directory. Make sure that the temporary directory specified has enough space for the files being transferred.

```
your_localhost% sup rsync -T /tmp foobar pfe1:
```

scp (man page)

```
your_localhost% sup scp foobar pfe1:/tmp/c_foobar
```

sftp (man page)

```
your_localhost% sup sftp pfe1
```

File Monitoring Command

test (man page)

```
your_localhost% sup ssh pfe1 test -f /tmp/c_foobar
```

Job Monitoring Command

qstat (man page available on Pleiades and Columbia)

```
your_localhost% sup ssh pfe1 qstat @pbspl1
```

SUP Expected Output

The following sequence shows the expected output for the command:

```
your_localhost% sup scp foobar pfe1:/tmp/c_foobar
```

for a user who has never used the SUP before.

The conditions under which each sub-sequence will be seen are indicated next to each header. Most of the items will only be seen once or during key generation. A second invocation will only show the command output portion.

1. **Host key verification** (seen once per client host)

```
No host key found for sup-key.nas.nasa.gov
...continue if fingerprint is
1b:9a:82:2b:b9:b0:7d:e5:08:50:1d:e8:14:76:a2:2e
The authenticity of host 'sup-key.nas.nasa.gov (129.99.242.7)'
can't be established.
RSA key fingerprint is
1b:9a:82:2b:b9:b0:7d:e5:08:50:1d:e8:14:76:a2:2e.
Are you sure you want to continue connecting (yes/no)? yes
No host key found for sup.nas.nasa.gov
...continue if fingerprint is
52:f3:61:9b:9c:73:79:4d:22:cb:f3:cd:9a:29:4e:fe
The authenticity of host 'sup.nas.nasa.gov (129.99.242.6)'
can't be established.
RSA key fingerprint is
52:f3:61:9b:9c:73:79:4d:22:cb:f3:cd:9a:29:4e:fe.
Are you sure you want to continue connecting (yes/no)? yes
```

2. **Identity creation** (seen during key generation if no identity available)

```
Cannot find identity /home/user/.ssh/id_rsa
...do you wish to generate it? (y/n) y
Generating public/private rsa key pair.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/user/.ssh/id_rsa.
Your public key has been saved in /home/user/.ssh/id_rsa.pub.
The key fingerprint is:
a3:cf:e5:50:12:6f:14:b1:21:59:19:a8:33:aa:77:40 user@host
```

3. **Identity addition to agent** (seen during key generation)

```
Adding identity /home/user/.ssh/id_rsa to agent
Enter passphrase for /home/user/.ssh/id_rsa:
Identity added: /home/user/.ssh/id_rsa
(/home/user/.ssh/id_rsa)
```

4. **Identity initialization** (seen once per identity)

```
Initializing identity on sup-key.nas.nasa.gov (provide login
information)
Password:
Enter PASSCODE:
```

```
Key a3:cf:e5:50:12:6f:14:b1:21:59:19:a8:33:aa:77:40 uploaded
successfully
```

5. **SUP key generation** (seen when no valid SUP keys available)

   ```
   Generating key on sup.nas.nasa.gov (provide login information)
   Password:
   Enter PASSCODE:
   ```

6. **Client upgrade** (seen during key generation when new client available)

   ```
   A newer version of the client is available (0.39 vs. 0.37)
   ...do you wish to replace the current version? (y/n) y
   ```

7. **Command output** (always seen)

   ```
   foobar 100% 5 0.0KB/s 00:00
   ```

SUP Troubleshooting

The following error messages may be encountered during your SUP client usage. Note that the "-v" option can be given to the SUP client to output additional debugging information.

- "WARNING: Your password has expired"

  This message indicates that your current password has expired and must be changed. To change your password, you must log in to an LDAP host (for example, Lou) through the SFEs and change your LDAP password. This change will be automatically propagated to the SUP within a few minutes.

- "Permission denied (*~/.meshrc* not found)"

  This message indicates that you have not created a *.meshrc* file in your home directory on the target host. SUP commands must be authorized for execution on each target host.

- "Permission denied (unauthorized command)"

  This message indicates that you have attempted an operation that is not currently authorized by the SUP. Check that the command line is valid and that the attempted command is one of the authorized commands. Certain options to authorized commands may also be disallowed, but these should never be needed in standard usage scenarios.

- Permission denied during file access (various forms)

These messages indicate that you attempted to read or write a file for which such access is not allowed. The most common cause is forgetting to authorize directories for writes. Reads and writes of *~/.** are never permitted.

File Staging through DMZ File Servers

The NAS DMZ (Demilitarized Zone) file transfer servers, dmzfs1.nas.nasa.gov and dmzfs2.nas.nasa.gov, are designed to help facilitate file transfers into and out of the NAS enclave. All Lou users have an account on the DMZ file servers.

Design

- Each DMZ server is independent; they do not share filesystems or data.
- The DMZs do not support RSA SecurID authentication, so, the RSA key fob is not needed, and setting up SSH passthrough is not required. Instead, a password or public/private key pair should be used for authentication
- SCP and bbFTP are supported file transfer protocols.

Setup

To set up public key authentication for the DMZs, copy the public key, which you have likely already created on your local host, to the *authorized_keys* file of dmzfs1 and/or dmzfs2:

```
localhost% scp ~/.ssh/id_rsa.pub nas_username@dmzfs1.nas.nasa.gov:~/.ssh
localhost% ssh nas_username@dmzfs1.nas.nasa.gov
dmzfs1% cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

- Files should be pushed to or pulled from the DMZs.
- Unattended file transfers via the DMZs can be done with public key authentication. Files generated inside the NAS HECC Enclave can be pushed to the DMZ file servers under script control (but *not through PBS jobs*). Likewise, remote systems can automatically push files to the DMZ file servers. Then, scripts operating on Pleiades or Columbia can periodically check for file availability on the DMZ file servers, and when available, will pull the file into Pleiades or Columbia.

Restrictions

- The user environments are jailed; executable commands are minimal.

- Outbound connections are not allowed. File transfers via the DMZ file servers using commands such as *scp* or *bbftp* must be initiated from your local host or NAS systems (such as Pleiades, Columbia, Lou) *not* dmzfs1 or dmzfs2.

- Storage space is limited (users share 2.5TB), and files are meant to be stored for very short durations. Every hour, files older than 24 hours are automatically removed.

Examples

The following examples assume that: a) You want to push a file to dmzfs1 from your local host and pull the file from pfe1; b) You have not set up public key authentication for the DMZs. Thus, password authentication is used.

- Using *scp,* first copy the file to the DMZ:

  ```
  localhost% scp foo dmzfs1.nas.nasa.gov:
  Password: <-- type in your lou password
  foo 100% 764 0.8KB/s 00:00
  ```

  If your NAS username and local username are different:

  ```
  localhost% scp foo nas_username@dmzfs1.nas.nasa.gov:
  Password:   <-- type in your lou password
  foo                 100%  764     0.8KB/s   00:00
  ```

  then, you can pull the file from the DMZ:

  ```
  pfe1% scp dmzfs1:foo .
  Password:  <-- type in your lou password
  foo                 100%  764     0.8KB/s   00:00
  ```
- Using *bbftp,* first copy the file to the DMZ:

  ```
  localhost% bbftp -s -e 'put foo' dmzfs1.nas.nasa.gov
  Password: <-- type in your lou password
  foo 100% 764 0.8KB/s 00:00
  ```

  If your NAS username and local username are different:

  ```
  localhost% bbftp -s -u nas_username -e 'put foo' dmzfs1.nas.nasa.gov
  Password:   <-- type in your lou password
  put foo OK
  ```

  then, you can pull the file from the DMZ:

  ```
  pfe1% bbftp -s -e 'get foo' dmzfs1
  Password:   <-- type in your lou password
  get foo OK
  ```

  See the article on bbftp for more instructions.

bbftp

DRAFT

This article is being reviewed for completeness and technical accuracy.

When and Why to use bbFTP

If your data is being transferred to or from a NAS system over the wide area network, scp will almost always be the limiting factor, due to the static TCP windowing defined in the OpenSSH (versions older than 4.7) source code. The Bandwidth Delay Product (BDP) states that the bandwidth of the pipe multiplied by the latency gives the optimal window size for data transfer. With the window size statically defined for lower-speed networks, scp can never properly utilize the bandwidth available. bbFTP has dynamically adjustable window sizes (up to the maximum allowed by the system) and can also transmit multiple simultaneous streams of data. We have found that this application provides the best mechanism for making use of the bandwidth available between two sites.

Things to check:

- Are you using scp to transfer files?
- Are you transferring files to an offsite location? (outside NAS or NASA Ames)
- Is the average delay between sites larger than 30 ms?
- Is the data being transferred in large files (1 GB+)?

If the answer to all of these is 'Yes', then the bbFTP application will improve data transfer rates. Please follow the guide below to get started.

Downloading bbFTP

bbFTP has been tested to work on many operating systems: Linux, IRIX, Solaris, BSD and MacOSX. Other systems may also be supported.

If you intend to intiaite bbFTP from your localhost, you will need to install the bbFTP client on your localhost. If you intend to initiate bbFTP from a NAS host, you will need to install the bbFTP server on your localhost.

- bbFTP for Linux, IRIX, Solaris, and BSD

  For Linux, IRIX, Solaris, and BSD systems, the bbFTP application can be downloaded from its distribution site IN2P3 in France. For your convenience, the latest version is available here.

  Download latest client version - bbftp-client-3.2.0 (GZ compressed file - 232 KB)

Download latest server version - bbftp-server-3.2.0 (GZ compressed file - 220 KB)

- bbFTP for MacOSX

  Download latest client version with fixes for MacOSX (binary - 252KB)

  Download latest server version with fixes for MacOSX (binary - 192KB)

Installing bbFTP

If you download a source code distribution, follow the instruction below to build and install bbFTP. This guide covers the client setup only. Installing the server version is similar.

```
your_localhost% tar -zxvf bbftp*
your_localhost% cd bbftp*/bbftpc  (or bbftp*/bbftpd  for the server version)
your_localhost% ./configure
your_localhost% make
your_localhost% make install (optional, requires root privileges to install)
```

By default, the application will install in /usr/local/bin. If you do not have admin privileges, you may skip the last step and copy the bbFTP binary to your home directory, or run it from the current location.

Using bbFTP

To write the version of bbftp and default values to standard output:

```
bbftp -v
```

For example:

```
pfe1% bbftp -v
bbftp version 3.2.0
Compiled with  :   default port 5021
                   compression with Zlib-1.2.3
                   encryption with OpenSSL 0.9.8a 11 Oct 2005
                   default ssh command = ssh -q
                   default ssh remote command = bbftpd -s
                   default number of tries = 5
                   default sendwinsize = 256 Kbytes
                   default recvwinsize = 256 Kbytes
                   default number of stream = 1
```

To request the execution of commands contained in the control file ControlFile or the ControlCommands using RemoteUsername on RemoteHost:

```
bbftp [Options] [-u RemoteUsername] -i ControlFile [RemoteHost]
bbftp [Options] [-u RemoteUsername] -e ControlCommands [RemoteHost]
```

Notice that -i or -e option are mandatory. The examples given in this article all use -e *ControlCommands*.

Available options are:

```
[-b (background)]
[-c (gzip compress)]
[-D[min:max] (Domain of Ephemeral Ports) ]
[-f errorfile]
[-E server command for ssh]
[-I ssh identity file]
[-L ssh command]
[-s (use ssh)]
[-S (use ssh in batch mode)]
[-m (special output for statistics)]
[-n (simulation mode: no data written)]
[-o outputfile]
[-p number of // streams]
[-q (for QBSS on)]
[-r number of tries ]
[-R .bbftprc filename]
[-t (timestamp)]
[-V (verbose)] will print out the transfer rate
[-w controlport]
[-W (print warning to stderr) ]
```

For more information about each option, see **man bbftp**. Those used in the examples will be briefly described.

Single stream vs multiple streams

- Single stream:

Using single stream is the easiest, but may not provide optimal performance.

In the examples below, bbFTP is run from the current working directory. If it was installed in a system path location, the "./" may be omitted.

The -s option says to use ssh to remotely start a bbftpd daemon. It usually starts the binary "bbftpd -s", but this can be changed througth the -E option.

The first command is to pull a file from a remotehost using *get* and the second command is to push a file to the remote host using *put*.

```
./bbftp -s -u remote_username -e 'get filename' remotehost
./bbftp -s -u remote_username -e 'put filename' remotehost
```

- Multiple streams:

For transfers between two NAS hosts, such as Pleiades and Lou, no more than 2 streams should be used.

For transfers between your site and NAS, more streams will probably help. In several tests, using 8 streams gave the best performance.

If there is little increase in the transfer rate from single stream to multiple streams, a lower number may be used. The value must be changed in both the control command (-e) and the server command (-E) so that the server listens for the same number of streams as the client requests.

In the examples below, -s is not used. Instead, -E 'bbftpd -s' is used to use ssh to remotely start a bbftpd daemon.

```
./bbftp -u remote_username -e 'setnbstream 8; get filename'
        -E 'bbftpd -s -m 8' remotehost
./bbftp -u remote_username -e 'setnbstream 8; put filename'
        -E 'bbftpd -s -m 8' remotehost
```

```
For formatting issue, each command above was broken into two lines.
In reality, it should be just one line.
```

- *File related commands*

  You may need to use the command 'cd' to change directory on the remotehost or 'lcd' to change directory on the host where bbftp is issued in order to 'get' or 'put' files from/to the directory you intend to use. For the rules, please see the man page of bbftp. Here are some examples:

  ```
  bbftp -s -u remote_username
        -e 'cd /u/username/abc; get filename' remotehost
  bbftp -s -u remote_username
        -e 'cd /u/username/abc; lcd def; put filename' remotehost
  ```

  ```
  For formatting issue, each command above was broken into two
  lines. In reality, it should be just one line.
  ```
- *Initiating bbftp from a host outside of NAS domain*

  If you want to initiate bbftp from a host that is not within the NAS domain to transfer files to/from a NAS host (not including dmzfs1 and dmzfs2), you must do the following:

  Set up SSH passthrough.

  In the *.ssh/config* file on your localhost, be sure to include entries with the **fully-qualified domain name**. For example:

  ```
  Host pfe1.nas.nasa.gov
  ```

```
ProxyCommand ssh sfe1.nas.nasa.gov /usr/local/bin/ssh-proxy pfe1.nas.nasa.gov
```

In the bbftp command line, use the **fully-qualified domain name** (ex: pfe1.nas.nasa.gov) of the NAS host. For example,

```
your_localhost% ./bbftp -s -u nas_username -e 'get filename'
pfe1.nas.nasa.gov
```

These two steps are needed due to the fact that bbftp uses 'gethostbyname' function to check a hostname for connection and then it uses ssh to connect to that hostname. Thus a fully-qualified domain name in the *./ssh/config* file is required. If the fully-qualified domain name cannot be found in *./ssh/config*, one will get the error:

```
BBFTP-ERROR-00061 : Error waiting MSG_LOGGED_STDIN message
```

For Pleiades, one has to use pfe[1-12].nas.nasa.gov or bridge[1-2].nas.nasa.gov. The front-end load balancer, **pfe.nas.nasa.gov**, does not work with bbftp. For example:

```
your_localhost% bbftp -s -u nas_username -e 'get filename' pfe.nas.nasa.gov
BBFTP-ERROR-00017 : Hostname no found (pfe.nas.nasa.gov)
```

On the other hand, for ssh or scp, one can use either the fully-qualified domain name above or the abbreviated name below:

```
Host pfe1
ProxyCommand ssh sfe1.nas.nasa.gov /usr/local/bin/ssh-proxy pfe1.nas.nasa.gov
```

- Specifying port range

Performance Tuning

To find the transfer rate, turn on the -V option.

Performance of bbFTP is affected by the number of streams and the TCP window sizes.

The TCP window size determines the amount of outstanding data a transmitting end-host can send on a particular connection before it gets acknowledgment back from the receiving end-host. For optimal performance, the window size should be set to the value of the Bandwidth Delay Product (i.e., the product of the bandwidth of the pipe and the latency).

bbFTP is compiled with a default send and receive TCP window size as can be seen with the -v option and can dynamically adjust the window size (up to the maximum allowed by the system) for better performance. However, a user can also choose a non-default send/recv window size (in KB). For example:

```
bbftp -e 'setrecvwinsize 1024; setsendwinsize 1024; put filename'
      -E 'bbftpd -s' remotehost
```

```
For formatting issue, the command above was broken into two lines.
In reality, it should be just one line.
```

For high-speed links where bbFTP is still not performing as well as expected, it may be due to a system windowing limitation. Most operating systems have the maximum window size set to a small value, such as 64 KB. As practice, NAS systems are set to a minimum of 512 KB.

If you are not gettting good performance, ask your local system administrator if performance tuning is necessary for your localhost.

The bbscp Script

DRAFT

This article is being reviewed for completeness and technical accuracy.

Introduction

The bbscp script is written in Perl by Greg Matthews at NAS. It is a bbftp wrapper which provides an scp-like command line interface; bbscp only encrypts usernames and passwords, it does *not* encrypt the data being transferred.

Downloading bbscp

If you plan to initiate bbscp on your localhost, you have to download bbscp version 1.0.6 (also attached at the end of this article) and download/install bbFTP client version 3.2.0 on your localhost.

The bbscp script has been installed on Pleiades (version 1.0.4), Columbia (version 1.0.4), and Lou (version 1.0.6).

Using bbSCP

Note that bbscp is just a client-side wrapper for bbftp, so, as with bbftp, you must use the *fully-qualified domain name* of the target host (for example, pfe1.nas.nasa.gov) if you are not within the NAS domain.

The bbSCP version 1.0.6 man page provides details on how to use it.

```
BBSCP(1)           User Contributed Perl Documentation              BBSCP(1)

NAME
       bbscp - bbftp wrapper, provides an scp-like commandline interface

SYNOPSIS
       bbscp [OPTIONS] [[user@]host1:]file_or_dir1 [...] [[user@]host2:]dir2

DESCRIPTION
       bbscp does unencrypted copies of files either from the localhost to a
       directory on a remote host, or from a remote host to a directory on
       the localhost (see the -N option for the only exception to this). It
       assembles the proper commandline for bbftp (designed and tested for
       bbftp version 3.2.0, see RESTRICTIONS) and then executes bbftp to
       perform the transfer(s).
       The "-s", "-p 2", and "-r 1" options for bbftp are set by default,
       along with the following options:

         setoption keepaccess
         setoption keepmode
         setoption nocreatedir

       The options -p and -r can be overridden on the commandline.

       Note the following limitations and capabilities in different transfer
       scenarios:

       copying from localhost to remote host
           - regular files
               bbftp will overwrite a pre-existing file of the same name on
               the remote host without asking for confirmation.

           - directories
               This script recursively transfers entire directories (only for
               local-to-remote transfers!).

           - symbolic links (see RESTRICTIONS)
               Symlinks on the localhost are treated just like the thing they
               point to, and are ignored if they point to something that
               doesn't exist.

       copying from remote host to localhost
           - regular files
               bbftp will overwrite a pre-existing file of the same name on
               the localhost without asking for confirmation.

           - directories
               There is no way at this time to transfer entire directories
               from a remote host to the localhost.

           - symbolic links (see RESTRICTIONS)
               Symlinks on the remote host are treated just like the thing
               they point to (which means they are ignored if they point to
               a directory or to something that doesn't exist).

OUTPUT
       The default output mode of the script displays "OK" or "FAILURE" for
       each of the transfer operations that bbftp performs. This display
       occurs after bbftp has finished running, so it may be delayed for
       some time depending on the duration of the transfer(s).

       The script switches to more verbose output if the user provides 1 or
       more of the verbose output commandline options (-l, -t, -V, and -W).

OPTIONS
         -B    name/location of bbftp executable. default is "bbftp"

         -d    dry-run. script performs its duty but does not actually
               execute bbftp. the bbftp commandline is printed, along
               with the contents of the bbftp control-file
```

```
          -h    minimal help text

          -k    keep bbftp command file that this script creates

          -l    long-winded (extra verbose) output from bbftp. uses
                undocumented bbftp option (-d)

          -N    transfer a single file and rename it at the destination.
                both local-to-remote and remote-to-local transfer is
                supported. see RESTRICTIONS

          -v    version of this script

          -X    set the size of the TCP send window (in kilobytes). default
                is the bbftp default size

          -Y    set the size of the TCP receive window (in kilobytes). default
                is the bbftp default size

          -z    suppress the security disclaimer

        bbftp options that can be specified on the commandline of this script:

          -D[min_port:max_port]   (e.g. "-D", "-D40000:40100")

          -E <Server command to run>

          -L <SSH command>

          -p <number of parallel streams>

          -R <bbftprc file>

          -r <number of tries>

          -t

          -V
          -W

RESTRICTIONS
        Version of bbftp
          It's very important to use bbftp version 3.2.0 with bbscp --
          there's at least 1 known issue with using bbftp 3.1.0.

        Possible shell issues
          bash and tcsh interpret commandline text in different ways, so you
          may need to use quotes or other delimiters to use bbscp. In
          particular, bash and tcsh are known to handle wildcards differently.

        Wildcards
          If the -N option is not in use, wildcards can be used in remote host
          file specifications, but only for the names of files, not for
          directories. So, for example, "user@host:/tmp/file*" is acceptable,
          but "user@host:/tm*/file*" is not.
```

```
       Symbolic links
         Symlinks are not bbftp's strong suit -- if you wish to transfer a
         collection of files that includes symlinks it is highly recommended
         that you first make a tar-file and then transfer the tar-file.

       Use of -N option
         Wildcards are not supported in remote host file specifications w/ -N.

         If the destination is a symlink it will be overwritten, regardless of
         what that symlink points to.

EXAMPLES
    Note: these examples have been tested with bash, changes may be needed for
    them to work in tcsh (see RESTRICTIONS).

    local file to remote directory (username must be the same on both machines)
         bbscp /u/username/data/file1 machine:target_dir

    local file to remote file w/ different name
         bbscp -N /u/username/data/file1 machine:file89

    multiple local files to remote directory
         bbscp /u/username1/data/*file username2@machine:/tmp

    local directory to remote home directory
         bbscp /u/username1/data username2@machine:

    remote file to local directory
         bbscp username1@machine:data/file5 /u/username2/source_dir

    remote file to local file w/ different name
         bbscp -N username1@machine:data/file5 /u/username2/source_dir/file93

    multiple remote files to local directory
         bbscp -V username1@machine:/u/username1/data/file* /tmp

    multiple remote files to local directory
         bbscp -V username1@machine:file1.txt username1@machine:stuff.dat /tmp

AUTHOR
      Greg Matthews gregory.matthews@nasa.gov

perl v5.8.8       2010-12-10    BBSCP(1)
```

Performance Tuning

To find the transfer rate, turn on -V option.

Like bbftp, the number of streams and TCP send/recv window sizes affect performance. Users can set the number of streams through the -p option. Starting with bbscp version 1.0.6, *the default is 2 streams***.** To set the window sizes in KB, use the *-X* option for send window and *-Y* for receive window. The efault is the bbftp default send/recv window size.

Test and Verification

bbscp man page

DRAFT

This article is being reviewed for completeness and technical accuracy.

The man page for bbscp as seen on Lou.

Using bbscp for Test and Verification

DRAFT

This article is being reviewed for completeness and technical accuracy.

The following examples provide test and verification data and sample commands for using bbscp between two hosts (crow & cfe3.nas.nasa.gov or dmzfs1.nas.nasa.gov).

Straight file transfer

This example demonstrates the tranfer of a file named *100mb*.

```
crow% bbscp -V 100mb user@cfe3.nas.nasa.gov:/nobackup1/user/

/home/user/bin/bbscp: will run commandline:
   bbftp -s -r 1 -V -p 8 -u user -i /tmp/bbscp.lKCrSUg cfe3.nas.nasa.gov

/home/user/bin/bbscp: begin output of bbftp:

-----------------------------------------------------------------
WARNING! This is a US Government computer. This system is for
.....
-----------------------------------------------------------------
Authenticated with partial success.

Plugin authentication

Enter PASSCODE:

>> COMMAND : setoption keepaccess
<< OK
>> COMMAND : setoption keepmode
<<  OK
>> COMMAND : setoption nocreatedir
<<  OK
>> COMMAND : put 100mb /nobackup1/user/100mb
<<  OK
104857600 bytes send in 5.43 secs (1.89e+04 Kbytes/sec or 147 Mbits/s)

/home/user/bin/bbscp: end output of bbftp
```

Renaming file at destination

Transfer a single file (named *100mb*) and rename it (to *crow-100mb*) at the destination; both local-to-remote and remote-to-local transfer is supported.

```
crow% bbscp -V -N 100mb user@cfe3.nas.nasa.gov:/nobackup1/user/crow-100mb

/home/user/bin/bbscp: will run commandline:
                      bbftp -s -r 1 -V -p 8 -u user -i

/tmp/bbscp.5eUBcTX cfe3.nas.nasa.gov

/home/user/bin/bbscp: begin output of bbftp:

----------------------------------------------------------------

WARNING! This is a US Government computer. This system is for
.....
----------------------------------------------------------------

Authenticated with partial success.

Plugin authentication

Enter PASSCODE:

>> COMMAND : setoption keepaccess
<< OK
>> COMMAND : setoption keepmode
<< OK
>> COMMAND : setoption nocreatedir
<< OK
>> COMMAND : put 100mb /nobackup1/user/crow-100mb
<< OK
104857600 bytes send in 5.3 secs (1.93e+04 Kbytes/sec or 151 Mbits/s)

/home/user/bin/bbscp: end output of bbftp
```

Adjusting the TCP window size

This example demonstrates the use of -X and -Y options to set the TCP window size (available in bbscp Version 1.0.2 and above).

```
crow% ./bbscp -V -N -X 2000 -Y 2000 1gig.dat user@dmzfs1.nas.nasa.gov:/home/user/garbag

bbscp: will run commandline:
       bbftp -s -r 1 -V -p 8 -u kfreeman
             -i /tmp/bbscp.SNxL5RT dmzfs1.nas.nasa.gov

bbscp: begin output of bbftp:

user@dmzfs1.nas.nasa.gov's password:

>> COMMAND : setoption keepaccess
<< OK
>> COMMAND : setoption keepmode
```

```
<< OK
>> COMMAND : setoption nocreatedir
<< OK
>> COMMAND : setsendwinsize 2000
<< OK
>> COMMAND : setrecvwinsize 2000
<< OK
>> COMMAND : put 1gig.dat /home/kfreeman/garbage.dat
<< OK

1109393408 bytes send in 34.6 secs (3.13e+04 Kbytes/sec or 244 Mbits/s)

bbscp: end output of bbftp
```

Dry run/debugging

This example demonstrates the use of the -d option for dry run. In this case, the bbscp script performs its duty but does not actually execute bbFTP. The bbFTP command line is printed, along with the contents of the bbFTP control-file.

```
cfe3.user% bbscp -d -V -N one-gig user@crow.eos.nasa.gov:/home/user/data/cfe3-one-gig
/usr/local/bin/bbscp: would have run commandline:
                      bbftp -s -r 1 -V -p 8 -u user
                      -i /tmp/bbscp.4PZYIuL crow.eos.nasa.gov

/usr/local/bin/bbscp: bbftp control-file (/tmp/bbscp.4PZYIuL) looks like:

setoption keepaccess
setoption keepmode
setoption nocreatedir
put one-gig /home/user/data/cfe3-one-gig
```

Using the SUP Virtual File System

DRAFT

This article is being reviewed for completeness and technical accuracy.

Introduction

The SUP client includes a virtual file system (VFS) capability that allows files across all SUP connected resources to be accessed using standard file system commands. For example, the command:

```
ls /sup/pfe1/tmp
```

would list the files in /tmp on *pfe1*. The command:

```
cp foobar /sup/pfe1/tmp
```

would copy the file "foobar" from the current directory on the local host to /tmp on *pfe1*.

The set of supported commands includes cat, cd, chgrp, chmod, chown, cmp, cp, df, diff, du, file, grep, head, less, ln, ls, mkdir, more, mv, pwd, rm, rmdir, tail, tee, test, touch, and wc. Note that this functionality is not a true file system since only these commands are supported and only when used from within a shell. Unlike more general approaches such as FUSE, however, the SUP capability is completely portable and can be enabled with no additional privileges or software.

Commands through the VFS functionality can act on any combination of local and remote files, where remote files are prefixed with "/sup/hostname". For example, the command:

```
cat /sup/pfe1/tmp/rfile ~/lfile
```

would print the file "rfile" in /tmp on *pfe1* as well as the file "lfile" in the user's home directory on the local host to the terminal. Any number of hosts can be included in any command. For example, the command:

```
diff /sup/pfe1/tmp/cfe_file /sup/pfe/tmp/pfe_file
```

would show the differences between the file "cfe_file" in /tmp on *pfe1* and the file "pfe_file" in /tmp on pfe. The client determines if any remote access is needed based on the path(s) given. If not, it will execute the command locally as given as rapidly as possible. Fully local commands also support all options with the exception of options of the form "-f value" (i.e. single-dash options that take values).

VFS Activation

- Requirements

  Currently, SUP VFS functionality is **only supported for bash**, but csh support is planned for the future. This functionality **requires Perl version 5.8.5** (note that this is more recent than version 5.6.1 required by the basic client functionality). It also **requires the standard Unix utilities cat, column, false, sort, and true** and has been tested successfully on Linux, OS X, and Windows under Cygwin and coLinux. Note that users of Windows under Cygwin may need to install the coreutils and util-linux packages to obtain these utilities.
- Activation/Deactivation

  1. Install the SUP client if you have not already done so
  2. Activate VFS functionality in a bash shell

     ```
     eval `sup -s bash`
     ```

     This will load aliases and functions used to intercept specific commands and replace them with commands through the SUP client that perform the actions requested.

  3. Deactivate VFS functionality in a bash shell whenever desired

     ```
     eval `sup -r bash`
     ```
- Command-line Options

  The behavior of the virtual file system can be modified using various options at the time it is activated.

  - -m /newroot

    Change the root of the virtual file system from its "/sup" default to "/newroot".

  - -ocmd=opts

    Specify default options for a given command since the VFS functionality overrides any existing aliases for its supported set of commands.

  - -t transport

    Change the file transport from its "sftp" default to "transport". Currently, the only additional transport available is "bbftp". Note that using bbftp as the transport may slow down certain operations on small files as bbftp has higher startup overhead.

  - -u user

Specify NAS user name. Note that this option **is required if your local user name differs from your NAS user name**.

For example, the following invocation activates the client virtual file system using bbftp as the transport mechanism, "nasuser" as the user and adds colorization of local file listings using the Linux ls "--color=always" option.

```
eval `sup -s bash -t bbftp -u nasuser -ols=--color=always`
```

VFS Caveats

The VFS functionality is still somewhat experimental. In general, it works for the most common usage scenarios with some caveats. In particular:

- "Whole file" commands (i.e. commands that must process the entire file), including cat, cmp, diff, grep, wc (and currently more/less due to implementation) retrieve files first before processing for efficiency. Thus, these commands should not be executed on very large files.
- There is a conflict between commands that take piped input and the custom globbing of the client, thus these commands have portions of globbing support disabled. These commands are grep, head, less, more, tail, tee, and wc. In these cases, globbing will work for absolute prefixes, but not relative. For example, "grep foo /sup/pfe1/tmp/*" will work, but "cd /sup/pfe1/tmp; grep foo *" will not.

- Redirection to/from remote files doesn't work. The same effect can be achieved using cat and tee (e.g. "grep localhost a" would become "cat /sup/pfe1/etc/hosts |grep localhost |tee a >/dev/null"). Redirection still works normally for local files.
- The directories "/sup" and "/sup/hostname" show up in neither completions nor ls, so you must know they exist.
- The first time a command is run involving a particular host, a SFTP connection is created to that host. When running "ps", it may appear as if a zombie client process is running.
- Commands may hang the first time after switching networks (e.g. with a laptop). If this happens, hit Control-c and it will work the next time.

VFS Commands

Currently supported commands and their currently supported options are below. Unsupported options will simply be ignored except where noted. All commands are still subject to SUP authorizations, thus something that cannot be executed or written normally through the SUP cannot be executed or written through this functionality either.

- **cat (no options)**
- **cd (no options)**

Note that when changing to remote directories, cd only changes $PWD so to make changes visible, the working directory (i.e. \w in bash) must be in your prompt. For example, the following prompt:

```
export PS1="\h:\w> "
```

would display the current host name followed by the current working directory.

- **chgrp (no options)**

  Groups may be specified either by number or by name. Names will be resolved on the remote host.

- **chmod (no options)**

  Modes must be specified numerically (e.g. 0700). Symbolic modes, such as a+rX, are not currently supported.

- **chown (no options)**

  Users and groups may be specified either by number or by name. Names will be resolved on the remote host.

- **cmp (all options)**
- **cp [-r]**

  Note that copies between two remote hosts transfer files to the local host first since the SUP does not allow third party transfers. Thus, very large file transfers between remote systems should be achieved using an alternate approach.

- **df [-i]**

  Note that 1024-byte blocks are used.

- **diff (all options)**
- **du [-a] [-b] [-s]**

  Note that 1024-byte blocks are used.

- **file (all options)**

- **grep (all options)**

- **head [-number]**

  Note that head does not support the form "-n number", thus, for example, to display the first 5 lines of a file, use "-5" and not "-n 5".

- **less (all options)**
- **ln [-s]**

  Note that hard links are not supported. Links from remote files to local files (e.g. ln -s /sup/pfe1/foo /foo) will be dereferenced during certain operations (e.g. cat /foo will cat /sup/pfe1/foo).

- **ls [-1] [-d] [-l]**

  For efficiency purposes, ls behaves slightly differently for remote commands than for local. In particular "ls -l" will not show links by default and will show what is actually linked instead of the link itself. Link details can be obtained using the "-d" option (e.g. ls -ld *).

  Also for efficiency, ls processes remote files before local files, so output ordering may be changed when remote and local files are interleaved on the ls command line. For example, "ls /foo/sup/pfe1/bar" would show /sup/pfe1 first, then /foo, then /bar.
- **mkdir (no options)**
- **more (all options)**
- **mv (no options)**
- **pwd (no options)**
- **rm [-r]**
- **rmdir (no options)**
- **tail [-number]**

  Note that tail does not support the form "-n number", thus, for example, to display the last 5 lines of a file, use "-5" and not "-n 5".

- **tee [-a]**
- **test [-b] [-c] [-d] [-e] [-f] [-g] [-h] [-k] [-L] [-p] [-r] [-s] [-S] [-u] [-w]**

  Note that compound and string tests are not supported. Compound and string tests can be achieved using multiple test commands separated by shell compound operators. For example,

  ```
  test -f /sup/pfe1/foo -a "abc" != "123"
  ```

  would become

```
test -f /sup/pfe1/foo && test "abc" != "123"
```

Alternatively, the "actual" test command can be executed through the SUP:

```
sup ssh pfe1 test -f /foo -a "abc" != "123"
```

- **touch (no options)**
- **wc (all options)**

Using the SUP without the SUP Client

DRAFT

This article is being reviewed for completeness and technical accuracy.

Introduction

The SUP client is the recommended approach to using the SUP. The client requires Perl, however, thus may not be suitable for all purposes. The only software actually required to use the SUP is SSH. This page details the manual steps required to use the SUP with only SSH. Users should still review the client instructions for a full overview of the SUP.

SUP Manual Usage Summary

The steps below demonstrate how to get up and running with the SUP without the client using a bbftp transfer to cfe1 as an example. Consult the link in each step for full details (or simply read this page to completion).

1. Initialize a long-term key on sup-key.nas.nasa.gov (one time)

   ```
   ssh -x -oPubkeyAuthentication=no sup-key.nas.nasa.gov \
       mesh-keygen --init <~/.ssh/authorized_keys
   ```
2. Generate a SUP key (one time per week)

   ```
   eval `ssh-agent`
   ssh-add ~/.ssh/id_rsa
   ssh -A -oPubkeyAuthentication=no sup.nas.nasa.gov \
       mesh-keygen |tee ~/.ssh/supkey`
   ssh-agent -k
   ```
3. Authorize host for SUP operations (one time per host)

   ```
   ssh cfe1
   touch ~/.meshrc
   ```
4. Authorize directories for writes (one or more times per host)

   ```
   ssh cfe1
   echo /tmp >>~/.meshrc
   ```
5. Prepare the SUP key for use (one time per session)

   ```
   eval `ssh-agent`
   ssh-add -t 1w ~/.ssh/supkey
   ```
6. Execute command (each time)

   ```
   bbftp -L "ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q" \
       -e "put /foo/bar /tmp/c_foobar" cfe1.nas.nasa.gov
   ```

7. Troubleshoot problems (as needed)

SUP Key Generation

1. **On the very first use only**, invoke the "mesh-keygen" command with the "--init" option on sup-key.nas.nasa.gov to upload an SSH authorized_keys file (used *only* during key generation and revocation). An authorized_keys file contains one or more SSH public keys that allow the corresponding SSH private keys to be used for authentication to a system. The uploaded authorized_keys file can be an existing file (such as your ~/.ssh/authorized_keys file from any host) or one created specifically for this purpose using a new SSH key pair generated with ssh-keygen. The public keys in this file must be in OpenSSH format (i.e. *not* the format of the commercial SSH version used on the Secure Front-Ends [SFEs]) and must not contain any forced commands (i.e. "command="). For example, to upload an existing authorized_keys file, the following can be invoked:

   ```
   ssh -x -oPubkeyAuthentication=no sup-key.nas.nasa.gov \
       mesh-keygen --init <~/.ssh/authorized_keys
   ```

   You will be prompted to authenticate using both a password (originally your Lou password) and securID passcode (PIN + tokencode).

   Users who have never connected to sup-key.nas.nasa.gov before may need to add a "-oStrictHostKeyChecking=ask" option to the scp command line. (RSA key fingerprint of sup-key.nas.nasa.gov is 1b:9a:82:2b:b9:b0:7d:e5:08:50:1d:e8:14:76:a2:2e)

   Note that this is on sup-key *only* and that you must use the "-oPubkeyAuthentication=no" option as shown. Users outside NAS may need to add an appropriate SSH option to set their login name, such as "-l username".
2. Start an SSH agent (or use one currently running):

   ```
   eval `ssh-agent -s` (if your shell is sh/bash)
   ```

   or

   ```
   eval `ssh-agent -c` (if your shell is csh/tcsh)
   ```

3. Add a private key corresponding to one of the public keys in the authorized_keys file of step 1 to the agent (this is unnecessary if an agent is already running with the key loaded). For example:

   ```
   ssh-add ~/.ssh/id_rsa
   ```

4. Invoke the "mesh-keygen" command on sup.nas.nasa.gov. You will be prompted to authenticate using both password (originally your Lou password) and securID passcode (PIN + tokencode). After successful authentication, the mesh-keygen

command prints a SUP key to your terminal, which should be saved to a file in a directory that is readable only by you. This key can be saved to a file by cut-and-paste, redirecting standard output, or using the "tee" command. For example, to generate a key and redirect it into a file starting with ~/.ssh/supkey and labeled with the current time, the following can be invoked:

```
ssh -A -oPubkeyAuthentication=no sup.nas.nasa.gov \
    mesh-keygen |tee ~/.ssh/supkey.`date +%Y%m%d.%H%M`
```

Users who have never connected to sup.nas.nasa.gov before may need to add a "-oStrictHostKeyChecking=ask" option to the SSH command line. (RSA key fingerprint of sup.nas.nasa.gov is 52:f3:61:9b:9c:73:79:4d:22:cb:f3:cd:9a:29:4e:fe)

Note that you must use the "-oPubkeyAuthentication=no" option as shown. Users outside NAS may need to add an appropriate SSH option to set their login name, such as "-l username".

5. Protect your keys. In order to perform unattended operations, SUP keys cannot be encrypted, thus should always be protected with appropriate file system permissions (i.e. 400 or 600). Check the permissions of your key immediately after generation and modify if necessary. You are responsible for the privacy of your keys.

SUP Key Management

Each invocation of mesh-keygen creates a new SUP key that is valid for one week from the time of generation. Users may have multiple keys at once that all expire at different times. To facilitate the management of multiple SUP keys, the "mesh-keytime" and "mesh-keykill" commands are available.

Mesh-keytime

To determine the expiration time of a SUP key stored in a file "/key/file", the following can be invoked:

```
ssh -xi /key/file -oIdentitiesOnly=yes -oBatchMode=yes \
    sup.nas.nasa.gov mesh-keytime
```

The key fingerprint and expiration time will be printed to your terminal.

Mesh-keykill

To invalidate a specific SUP key stored in a file "/key/file" before its expiration time has passed, you must have an SSH agent running with the same key you use to generate SUP keys as described in steps 2 and 3 of the SUP Key Generation section. After which, the following can be invoked:

```
ssh -Axi /key/file -oIdentitiesOnly=yes -oBatchMode=yes \
            sup.nas.nasa.gov mesh-keykill
```

To invalidate all currently valid SUP keys, the following can be invoked:

```
ssh -Ax -oPubkeyAuthentication=no sup.nas.nasa.gov mesh-keykill --all
```

In this case, you will be prompted to authenticate using both password and securID passcode.

SUP Key Preparation

Currently, the only operations allowed with a SUP key are scp, sftp, bbftp, qstat, rsync, and test. For all operations, an SSH agent must be started with the SUP key loaded, which can be scripted as needed, because the key is unencrypted.

1. Start an SSH agent:

   ```
   eval `ssh-agent -s` (if your shell is sh/bash)
   ```

   or

   ```
   eval `ssh-agent -c` (if your shell is csh/tcsh)
   ```
2. Add a SUP key to the agent (this is the *only* key required to perform unattended SUP operations):

   ```
   ssh-add /key/file
   ```

   Since SUP keys have a lifetime of one week, the "-t" option may be used to automatically remove the key from the agent after a week has elapsed:

   ```
   ssh-add -t 1w /key/file
   ```

   The will prevent a buildup of keys in the agent, which can cause login failure as described in the SUP Troubleshooting section. Keys may be explicitly removed from the agent using the following:

   ```
   ssh-keygen -y -f /key/file >/key/file.pub
   ssh-add -d /key/file
   ```
3. Make sure agent forwarding and batch mode are enabled in your SSH client. The examples below include the appropriate options to enable agent forwarding ("-A") and batch mode ("-oBatchMode=yes").

SUP Commands

Examples of the use of each command that may be executed through the SUP are given below. Note that SUP commands must be authorized for execution on each target host and transfers to a given host must be authorized for writes.

- **bbftp** (man page)

```
bbftp -L "ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q" \
    -e "put /foo/bar /tmp/c_foobar" cfe1.nas.nasa.gov
```

Note that **you must use the fully-qualified domain name of the target host** (in this case, cfe1.nas.nasa.gov) if you are not within the NAS domain.

- **bbscp** (man page)

```
bbscp -L "ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q" \
    foobar cfe1.nas.nasa.gov:/tmp/c_foobar
```

Note that bbscp is just a client-side wrapper for bbftp, thus like bbftp, **you must use the fully-qualified domain name of the target host** (in this case, cfe1.nas.nasa.gov) if you are not within the NAS domain.

- **qstat** (man page available on Pleiades and Columbia))

```
ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q cfe1 qstat @pbs1
```

- **rsync** (man page)

```
rsync -e "ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q" \
    foobar cfe1:/tmp/c_foobar
```

Note that even if your home directory has been authorized for writes, **rsync transfers to your home directory will fail unless the "-T" or "--temp-dir" option is specified**. This is because rsync uses temporary files starting with "." during transfers, which cannot be written in your home directory. By specifying an alternate temporary directory that is authorized for writes, this problem can be avoided. For example, the following uses /tmp as the temporary directory when files are transferred to the home directory. Make sure that the temporary directory specified has enough space for the files being transferred.

```
rsync -T /tmp -e "ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q" \
    foobar cfe1:
```

- **scp** (man page)

  1. Create a file (for example, "supwrap") containing the following:

```
#!/bin/sh
 exec ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q $@
```

  2. Make the created file executable:

```
chmod 700 supwrap
```

  3. Initiate the transfer. For example:

```
scp -S ./supwrap foobar cfe1:/tmp/c_foobar
```

- **sftp** (man page)

  1. Create a file (for example, "supwrap") containing the following:

```
#!/bin/sh
```

```
            exec ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q $@
```

   Note that this file is identical to the one described for scp.
2. Make the created file executable:

   ```
   chmod 700 supwrap
   ```
3. Initiate the transfer. For example:

   ```
   sftp -S ./supwrap cfe1
   ```

- **test** (man page)

  ```
  ssh -Aqx -oBatchMode=yes sup.nas.nasa.gov ssh -q cfe1 test -f /tmp/c_foobar
  ```

SUP Troubleshooting

The following error messages may be encountered during SUP usage.

- "WARNING: Your password has expired"

  This message indicates that your current password has expired and must be changed. To change your password, you must log in to an LDAP host (e.g. Lou) through the SFEs and change your LDAP password. This change will be automatically propagated to the SUP within a few minutes.
- "Permission denied (~/.meshrc not found)"

  This message indicates that you have not created a .meshrc file in your home directory on the target host. SUP commands must be authorized for execution on each target host.
- "Permission denied (key expired)"

  SUP keys are only valid for one week from the time of generation. This message indicates that the SUP key used for authentication has expired and is no longer valid. You must generate a new SUP key or use a different SUP key before attempting another operation.
- "Permission denied (publickey,keyboard-interactive)"

  This message indicates that you have not provided the appropriate authentication credentials to the SUP. There may be several causes:

  - ♦ If you are generating a SUP key and also receive an "Error copying key..." message, you have not loaded a private key into your SSH agent corresponding to one of the public keys in the authorized_keys file uploaded to sup-key in steps 1-3 of the SUP Key Generation section. You can verify that the correct key is loaded by running "ssh-keygen -l -f uploaded_key_file" and "ssh-agent -l" and checking that the fingerprint of your uploaded key file has been loaded into your SSH agent.

    - ♦ If you have specified -oBatchMode=yes on the command line, a valid SUP key may not been loaded into your SSH agent. There may also be too many keys loaded into your agent. SSH tries each key in the agent sequentially, so a valid key may still fail if it was added to the agent after a number of invalid keys greater than or equal to the login attempt limit. Check the number of keys in the agent using "ssh -l". The agent may be cleared of keys using "ssh-add -D".

    - ♦ If you have specified -oPubkeyAuthentication=no, you have not provided a valid password and/or a valid securID passcode.

- "Permission denied (unauthorized command)"

  This message indicates that you have attempted an operation that is not currently authorized by the SUP. Check that the command line is valid and that the attempted command is one of the authorized commands. Certain options to authorized commands may also be disallowed, but these should never be needed in standard usage scenarios.
- Permission denied during file access (various forms)

  These messages indicate that you attempted to read or write a file for which such access is not allowed. The most common cause is forgetting to authorize directories for writes. Reads and writes of ~/.* are never permitted.

Using GPG to Encrypt Your Data

DRAFT

This article is being reviewed for completeness and technical accuracy.

Introduction

Inter-host file transfer (ex: scp, bbftp, ftp) is better protected when the files are encrypted. GPG (Gnu Privacy Guard) is an Open Source OpenPGP compatible encryption system that we recommend you to use for this purpose. GPG (version 1.4.2) has been installed on Pleiades, Columbia and Lou at /usr/bin/gpg. If you do not have GPG installed on the system(s) that you would like to use for file transfer, please check out the GPG web site. Information, HOWTOs, Guides, FAQs, etc., for GPG can be found at: http://www.gnupg.org

Choosing what cipher to use

We recommend using the cipher AES256, which uses a 256-bit AES key to encrypt the data. Information on AES can be found at:
http://csrc.nist.gov/CryptoToolkit/tkencryption.html

One can set the desired cipher in the following ways:

- add the following line to your ~/.gnupg/gpg.conf

  ```
  cipher-algo AES256
  ```

- add "--cipher-algo AES256" to override the default cipher CAST5 in the command line:

  For any of the following examples in the Simple Examples section, you can add "--cipher-algo AES256" to override the default cipher CAST5 if you chose to not add the "cipher-algo AES256" to your personal gpg.conf file.

Simple Examples

- creating an encrypted file:

  Both commands below are identical. They encrypt the file 'test.out' and produce the encrypted version in 'test.gpg'.

  ```
  gpg --output test.gpg --symmetric test.out
  gpg -o test.gpg -c test.out
  ```

  You will be prompted for a passphrase, which will be used later to decrypt the file.

- decrypting a file:

  The following command decrypts the file "test.gpg" and produces the file "test.out".

  ```
  gpg --output test.out -d test.gpg
  ```

  You will be prompted for the passphrase which you used to encrypt the file.

  If you don't use the "--output" option, output of the command goes to STDOUT.

  If you don't use any flags, it will decrypt to a file without the .gpg suffix. That is,

  ```
  gpg test.gpg
  ```

  results in the decrypted data in a file named "test".

Passphrase Selection

Your passphrase should have lots of entropy. We suggest that you include five words of 5-10 letters in size chosen at random with spaces and/or numbers embedded into words and special characters.

You need to be able to recall the passphrase that was used to encrypt the file.

Factors that Affect Encrypt/Decrypt Speed on NAS HECC Filesystems

We do not recommend using the --armour option for encrypting files that will be transferred to/from NAS HECC systems. This option is mainly to send binary data through email, not scp/bbftp/ftp, etc. The file size tends to be about 33% bigger than without this option and takes about 10-15% longer to encrypt the data.

The level of compression used when encrypting/decrypting affects the time required to complete the operation. There are three options for the compression algorithm: none, zip and zlib.

- *--compress-algo none* or *--compress-algo 0*
- *--compress-algo zip* or *--compress-algo 1*
- *--compress-algo zlib* or *--compress-algo 2*

For example,

```
gpg --output test.gpg --compress-algo zlib --symmetric test.out
```

If your data is not compressible, "--compress-algo 0" (aka none) gives you about a 50% performance increase compared to zip "--compress-algo 1" or zlib "--compress-algo 2".

If your data is highly compressible, choosing zlib or zip will not only give you a 20-50% speed increase, but also reduce the file size by upto 20x. For example, a 517MB highly compressible file was compressed to 30MB on Columbia.

zlib is not compatible with PGP 6.x, but neither is the cipher algorithm AES256. zlib is about 10% faster than zip on Columbia and compresses about 10% better than zip.

Random Benchmark Data

We tested the encryption/decription speed of three different files (1MB, 150MB, 517MB) on Columbia. The file used for the 1MB test was an rpm file, presumably already compressed, since the resultant file sizes for the none/zip/zlib were within 1% of each other. The 150MB file was an ISO, also assumed to be a compressed binary file for the same reasons. The 517MB file is a text file. These runs were performed on a CXFS filesystem and when many other users' jobs were running. Thus, the performance reported here is just for reference, not the best or worst performance you can expect.

- Using AES256 as the cipher algorithm without --armour:

```
1MB file took ~4 secs to encrypt.
150MB took ~35 secs to encrypt.
```

- Using AES256 as the cipher algorithm with --armour:

```
1MB file took ~5.5 secs to encrypt.
150MB took ~40 secs to encrypt.
```

- Using AES256 as the cipher algorithm without --armour, zlib compression:

```
150MB took ~33 secs to encrypt.
decrypt to file: ~28 secs
```

- Using AES256 as the cipher algorithm without --armour, zip compression:

```
150MB took ~36 secs to encrypt.
decrypt to file: ~31 secs
```

- Using AES256 as the cipher algorithm without --armour, no compression:

```
150MB took ~19 secs to encrypt.
decrypt to file: ~25 secs
```

- Using AES256 as the cipher algorithm without --armour, no compression:

```
517MB text file took ~49 secs, resultant filesize ~517MB
decrypt to file: ~75 secs
```

- Using AES256 as the cipher algorithm without --armour, zip compression:

```
517MB text file took ~38 secs, resultant filesize ~33MB
```

```
        decrypt to file: ~34 secs
```

- Using AES256 as the cipher algorithm without --armour, zlib compression:

```
        517MB text file took ~33 secs, resultant filesize ~30MB
        decrypt to file: ~34 secs
```

Checking File Integrity

DRAFT

This article is being reviewed for completeness and technical accuracy.

It is a good practice to check that your data are complete and accurate before and after a file transfer. A common way for checking data integrity is to compute a checksum of the data.

There are multiple algorithms and programs that one can use for computing a checksum. A good checksum algorithm will yield a different result with high probability when the data is accidentally corrupted. If the checksums obtained before and after the transfer match, the data is almost certainly not corrupted.

On NAS HECC systems, the following programs are available:

- *sum*

  computes a checksum using BSD sum or System V sum algorithm; also counts the number of blocks (1KB-block or 512B-block) in a file
- *cksum*

  computes a cyclic redundancy check (CRC) checksum; also counts the number of bytes in a file
- *md5sum*

  computes a 128-bit MD5 checksum which is represented by a 32-character hexadecimal number

For example,

```
%ls -l foo
-rw------- 1 username groupid  67358 Nov 15 11:49 foo

%sum foo
50063    66

%cksum foo
269056887 67358 foo

%md5sum foo
cfe0fc62607e9dc6ea0c231982316b75  foo
```

md5sum is more reliable than *sum* or *cksum* for detecting accidental file curruption, as the chances of accidentally having two files with identical MD5 checksum are extremely small. It is installed by default in most Unix, Linux, and Unix-like operating systems. Users are recommended to compute the *md5sum* of a file before and after the transfer.

The following example shows that the file *foo* is complete and accurate after the transfer based on its md5sum.

```
pfe1% md5sum foo
cfe0fc62607e9dc6ea0c231982316b75  foo

pfe1% scp foo local_username@your_localhost:

your_localhost%md5sum foo
cfe0fc62607e9dc6ea0c231982316b75  foo
```

See **sum, cksum, md5sum** man pages for more information.

File Transfers Tips

When transferring files to NAS systems, there may be some ways to improve your performance without modifying your system (see TPC Performance Tuning for WAN Transfers). Below are some quick and easy techniques to try that may improve your performance rates when transferring files remotely to or from NAS.

- Transfer files from the /nobackup file system, which is often faster than the locally mounted disks.
- If you are using SCP, try adding the "-C" option to enable compression:

  ```
  $ scp -C filename user@remotehost.com:
  ```

  This can sometimes double your performance rates.
- For SCP transfers, use a low process-overhead cipher like *arcfour*:

  ```
  $ scp -carcfour  filename user@remotehost.com:
  ```

  This can increase your rates by 5x, compared to older methods like *3des*.
- If you are transferring from Lou, make sure your file is online first. Use the following DFM commands to determine this:

  ```
  $ dmls -al filename    # show the status of your file.
  $ dmget filename       # retrieve your file from tape prior to transferring.
  ```

  Get the full list of DMF commands.
- Use the bridge nodes to transfer files instead using of the Pleiades and Columbia front ends. These hosts have 10-Gigabit interfaces and more memory to handle both multiple and large file transfers.
- If you are transferring many small files, try using the *tar* command to compress them into a single file prior to transfer. Copying one large file is faster than transferring many small files.
- For files larger than a gigabyte, we recommended using BBFTP software, which can achieve much faster rates than single-stream applications such as SCP or RSYNC.

If you continue experiencing slow transfers and want to work with a network engineer to help improve file transfers, please contact support@nas.nasa.gov.

Archiving Data

Archiving Data Overview

DRAFT

This article is being reviewed for completeness and technical accuracy.

The articles in this category provide basic information to assist you in archiving or retrieving files to/from the NAS long term storage systems, Lou1-2.

Mass Storage Systems Lou1-2 describes the disk and tape drives capacity, and how to find the storage system that is assigned to you.

Archiving Data in Compute Systems talks about backing up data from the home and /nobackup filesystems of Pleiades or Columbia to long term storage and what file transfer commands can be used for such tasks.

Quota Policy on Disk Space and Files provides information on the policy and soft and hard quota limits of disk space and files (inodes) of the home and /nobackup filesystems of Pleiades, Columbia and Lou.

Validating Files Transferred to Mass Storage by Size and Number describes a light way approach to validate the files transferred from Pleiades/Columbia to Lou.

Validating Files Transferred to Mass Storage with md5sum provides an example of checking the integrity (through the use of md5sum) of data transferred from Pleiades/Columbia to Lou.

Using the GNU Tar Utility provides examples of using *tar* to collect multiple files into a tar file for easier distribution or storage on Lou and to extract files out of a tar file later on if needed.

Using CPIO Utility describes the CPIO tool, which is similar to Tar, and provides a few examples on how to use it.

Data Migration Facility (DMF) Commands explains the usage of the DMF commands for managing files (such as listing, finding, migrating, un-migrating, or copying) stored on tapes.

Mass Storage Systems: Lou1 and Lou2

The NAS environment contains three mass storage systems, Lou1 and Lou2, to provide long-term data storage for users of our high-end computing systems. These storage systems are SGI Altix computers running the Linux operating system. The disk space for the three systems combined is about 290 terabytes (TB), which is split into filesystems ranging from 9-30 TB in size.

Which Lou System I Should Use?

Each user should be able to log into any of the Lou systems, but will only have storage space on the home filesystem of one of them. Follow the steps below to determine which system you should store data on.

1. Log in to either Lou1 or Lou2. For example:

   ```
   your_localhost% ssh nas_username@lou1.nas.nasa.gov
   ```
2. Type the command "mylou" to find out your mass storage host. For example:

   ```
   lou1% mylou

   Your Mass Storage host is lou2

   Store files there in your home directory, /u/your_nas_username
   ```

   Be aware that Lou1 and Lou2 do *not* share their home filesystems.
3. Use the home filesystem on Lou*X* (where *X* = 1 or 2) determined by the step above for your long-term storage. For example:

   ```
   pfe1% scp foo lou2:
   ```

Quota Limits On Lou

For Lou*X* (where *X* = 1 or 2) that is assigned to you, there are no disk quota limits on your home filesystem. On the other hand, there *are* limits on the number of files (inode):

- 250,000 inode soft limit (14-day grace period)
- 300,000 inode hard limit

See Policy on Disk Files Quotas for Lou for more information.

Data (Un)Migration Between Disk and Tapes

In addition to the disk space, Lou1 and 2 have a combined 64 LTO-4 tape drives. Each of the LTO-4 tapes holds 800 GB of uncompressed data. The total storage capacity is approximately 10 PB.

Data stored on Lou's home filesystems (disk) is automatically migrated to tapes whenever necessary to make space for more data. Two copies of your data are written to tape media in silos located in separate buildings.

Data migration (from disk to tape) and unmigration (from tape to disk) are managed by the SGI Data Migration Facility (DMF) and Tape Management Facility (TMF).

If you need some data that is only available on tapes, make sure to unmigrate the data from tape to your home filesystem on Lou before transferring it to other systems.

For more tips on how to use Lou more effectively, see Storage Best Practices.

Archiving Data in Compute Systems

DRAFT

This article is being reviewed for completeness and technical accuracy.

Archiving Data under Pleiades or Columbia Home Filesystems

Users' data under the home filesystem of any NAS HECC system (such as Pleiades or Columbia) are automatically backuped to Susan (an archive system accessible only to system administrators) under script control. The backup is done **daily**.

If you lost important data under a home filesystem and need to have it restored, please send a request to NAS Help Desk and provide the following:

- system name
- your username
- the directory and/or file name(s) that you wish to restore
- the date the data was last touched

Please note that disk space quota limits are imposed on the Pleiades and Columbia home filesystems. If you are over your disk space quota limits, reduce your usage by deleting unneeded files, copying important files to your home filesystem on Lou, or to storage space at your local site and then removing them from Pleiades or Columbia.

Archiving Data under Pleiades or Columbia Nobackup File Systems

Data under any /nobackup filesystem of any NAS HECC system (such as Pleiades /nobackupp[1,10-60], or Columbia /nobackup[21-24], /nobackup[1-2][a-j], etc.) are not backuped by NAS. Users are responsible to do their own backups.

Please note that disk space and inode quota limits are imposed on the Pleiades Lustre filesystems /nobackupp[1,10-60] and Columbia CXFS filesystems /nobackup1a-h and /nobackup2a-i. If you are over the soft quota limits, reduce the usage by removing some files and/or archive any important data to your home filesystem on Lou, or to the storage space at your local site.

In addition, when the overall usage of a /nobackup filesystem is near its capacity, files can be deleted by system administrators. Normally, the few top users are sent emails and asked to clean up their disk space.

See Policy on Disk Space Quotas of home and /nobackup File Systems for Pleiades/Columbia for more information.

What File Transfer Commands to Use

Note that the Columbia CXFS nobackup filesystems (e.g. /nobackup[1-2][a-i]) are mounted on the Mass Storage system (Lou1-2) that you are assigned to. As a result, you can log into LouX (where X is 1or 2) and copy the files from nobackup to your home directory on Lou. SGI has created a command called cxfscp, which is a tuned version of the cp command. You can copy files at up to 400MB/s sustained with cxfscp.

If you need to initiate the transfer from Columbia, then we recommend you use bbscp or bbftp if the file to be transferred does not need to be encrypted. If you need to encrypt the data, even within the HECC enclave, then scp should be used.

Transferring files from Pleiades to Mass Storage can be done with bbscp/bbftp or scp. Disk-to-disk copying to Mass Storage may be implemented in the near future.

Using the GNU Tar Utility

If you have multiple related files and directory trees that you would like to archive to Lou, it is better to collect them into a single archive file using the GNU tar utility. See Using the GNU Tar Utility to learn more.

Related Articles

- Pleiades Home Filesystem
- Columbia Home Filesystem
- Pleiades Lustre Filesystem
- Columbia CXFS Filesystems
- Local File Transfer Commands - cxfscp
- Remote File Transfer Commands

Quota Policy on Disk Space and Files

DRAFT

This article is being reviewed for completeness and technical accuracy.

Some NAS filesystems enforce quotas. Two kinds of quotas are supported: limits on the total disk space occupied by the user's files, and limits on how many files (represented by *inodes*) the user can store, irrespective of size. For quota purposes, directories count as files.

Further, there are two different limits: hard limits and soft limits. Hard limits cannot be exceeded, ever. Any attempt to use more than your hard limit will be refused with an error. Soft limits, on the other hand, can be exceeded temporarily. You can stay over your soft limit for a certain period of time (the grace period). If you remain over your soft limit for more than the grace period, the soft limit is enforced as a hard limit.

You will not be able to add or extend files until you get back under the soft limit. Usually, this means deleting unneeded files or copying important files elsewhere (perhaps the Lou archival storage system) and then removing them locally.

When you exceed your soft limit you will begin getting daily emails reminding you how long until the grace period expires. These are intended to be informative and not a demand to immediately remove files.

The following table summarizes the default space and inode quota limits enforced on Columbia, Pleiades and Lou.

Default Quotas on Disk Space and Files

	Columbia	**Pleiades**	**Lou**
$HOME	NFS	NFS	XFS
Space: soft	4 GB	8 GB	none
Space: hard	5 GB	10 GB	none
Inode: soft	none	none	250,000
Inode: hard	none	none	300,000
/nobackup	CXFS /nobackup1[a-g] /nobackup2[a-i]	Lustre /nobackupp[10-70]	N/A
Space: soft	200 GB	200 GB	N/A
Space: hard	400 GB	400 GB	N/A
Inode: soft	25,000	75,000	N/A

Inode: hard	50,000	100,000	N/A

Policy on Disk Space Quotas of home and /nobackup File Systems for Pleiades/Columbia

- If you exceed the soft quota, an email will be sent to inform you of your current disk space and how much of your grace period remains. It is expected that a user will exceed their soft limit as needed, however after 14 days, users who are still over their soft limit will have their batch queue access to Columbia/Pleiades disabled.

- If an account has been disabled for more than 14 days, then its Columbia/Pleiades data will be moved to the archive host, Lou, and kept there for 6 months before removal, unless the project lead requests to have the data moved to another account.

- If an account no longer has batch access to a system, then all data from that system should be moved off within 7 days (or sooner if the other projects need the space).

- If an account needs larger quota limits, please send an email justification to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager, Bill Thigpen, for approval.

Policy on Disk File Quotas for Lou

- There is no quota for file space on Lou1or Lou2 because the data is written to tape. There is a quota on the number of files you can have. Currently there is a soft limit of 250,000 files and a hard limit of 300,000 files.

- There is a 14 day grace period if soft limit is exceeded. An email will be sent to inform you of your current disk space and how much of your grace period remains. It is expected that a user will exceed their soft limit as needed, however after 14 days, users who are still over their soft limit will be unable to archive files until they have reduced their use to below the soft limit.

- If an account needs larger quota limits, please send an email justification to support@nas.nasa.gov. This will be reviewed by the HECC Deputy Project Manager, Bill Thigpen, for approval.

- The maximum size of a file moved to Lou should not exceed 30% of the size of your home filesystem on Lou. If you need to move files larger than this, please contact the NAS Help Desk (support@nas.nasa.gov) for assistance.

Validating Files Transferred to Mass Storage by Size and Number

DRAFT

This article is being reviewed for completeness and technical accuracy.

It is a good practice to check if files are copied correctly to Mass Storage. The simplest and most light weight method to check files sent to Lou1or Lou2 is to compare the number of files and the disk space between the original and the copy.

Example:

```
pfe1% du -sk mydatadir
353760  mydatadir
pfe1% find mydatadir | wc -l
51
pfe1% scp -rp mydatadir lou2:

lou2% du -sk --apparent-size mydatadir
353684  mydatadir
lou2% find mydatadir | wc -l
51
```

Here the sizes are close and the number of files matches exactly.

Note that in most cases the directory size will not match exactly. The *--apparent-size* option is necessary on the Lou systems because the data may reside on tape, not disk.

Data Migration Facility DMF Commands

At the NAS facility, certain SGI systems (primarily the Lou filesystems) support a virtual storage manager feature called DMF (Data Migration Facility). Its purpose is to allow users to keep an increased volume of data in the files under their home directories by migrating those files that are not currently in use to offline backing storage, thereby allowing active disk space to be available for active files.

Migrated files are retrieved to active disk when you attempt to read or write to them, and the whole process is substantially transparent, aside from a possible time lag while a file is being retrieved.

The process of migrating files by backing them up from disk to another medium is implemented using STK tape silos.

DMF Commands

User commands associated with DMF are below, followed by usage examples:

dmls	Directory listing showing file migration status.
dmget	Retrieve migrated files to disk.
dmput	Cause files to migrate to backup on tape.
dmfind	Find files under a directory hierarchy.
dmcopy	Copy all or part of offline files.
dmattr	List attributes of files.

The *dmls* command

The *dmls* command is much like the standard *ls* command for file or directory listing, with the addition of an option for displaying the DMF status of files. Actually, *dmls* is derived from the GNU ls command, so it has a few extra "bells-and-whistles" compared to standard *ls* command; for details, see the **dmls man page**.

Example showing the extra status field:

```
% ls -l
total 64
-rw-r-----    1 aeneuman madmag     14713 Mar  1 17:02 file1
-rw-r-----    1 aeneuman madmag     17564 Mar  1 17:02 file2

% dmls -l
total 33
-rw-r-----   1 aeneuman madmag      14713 Mar  1 17:02 (REG) file1
-rw-r-----   1 aeneuman madmag      17564 Mar  1 17:02 (REG) file2
```

DMF has several possible states for files. The first three of them are the most likely to appear:

`REG`	File is a "regular" file that exists only online, on active disk.
`DUL`	File is "dual-state": identical copies of its data exist online and offline. The online copy will persist if there is no demand for free space in its filesystem. When free space is needed, the online copy of its data is released, leaving just the offline copy; the file becomes "offline." If you make any change to a dual-state file, the offline copy becomes out of date and invalid, and the file is once again a "regular" file.
`OFL`	File's directory entry remains but its data blocks are located offline only.
`MIG`	File is in the process of migrating offline.
`UNM`	File is in the process of un-migrating back online.
`NMG`	File is nonmigratable.
`INV`	File's DMF state is unknown or "invalid," usually because it is in a filesystem that does not use DMF.

The *dmget* command

The dmget command explicitly requests an offline file to be retrieved to disk. This command is not strictly required in order to retrieve offline files; any program that tries to use the file will cause it to be retrieved first.

The following example shows that referencing an offline file retrieves it after a pause. Notice that *C* was "offlined." Using the command *file C* caused this file to be retrieved and its status changed from *OFL* to *DUL*.

```
% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 11:20 (OFL) C

% file B
[immediate response:]
B:              English text

% file C
[pause while file is retrieved, then:]
C:              English text

% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 11:20 (DUL) C
```

You can retrieve the file explicitly with the command *dmget*. For example:

```
% dmls -l
total 220
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 11:20 (OFL) C

% dmget C ; echo Done
[pause, then:]
Done
```

One motivation for retrieving files explicitly with *dmget* is that you may be able to save time. If you are working with several files that have been migrated at the same time and reside on the same offline tape cartridge, a *dmget* command that names all the files would be able to retrieve them all in a single mount of the tape.

By contrast, simply using the files, one after another, would cause the tape robot to fetch the tape cartridge, retrieve the first file and put away the tape, then go back and fetch the cartridge, retrieve the second file, put away the tape, and so on.

Another good reason for retrieving files explicitly with *dmget* is that you control when the retrieval takes place. For example, suppose you have a large, multiprocessor application that is going to read a currently migrated data file. You would prefer the retrieval process to take place with just your shell running, not when your main application has spawned several dozen processes on several dozen nodes, with all of them having to sit idle waiting for the file to be retrieved.

See the example under dmfind for an efficient method to recall all files in a directory and its subdirectories.

The *dmput* command

In filesystems that are under DMF automated space management, large files that have not been used for a while will be migrated offline automatically. The definitions of "large" and "a while" are established on each system by its system administrator.

You can invoke migration explicitly with the *dmput* command. Useful options are:

-r	Causes DMF to release a file's online disk space immediately, giving it Offline status. Otherwise, the file would be in Dual status until its disk space was needed for other files.
-w	Causes the *dmput* command to wait until migration has completed before returning control. Otherwise the *dmput* command returns immediately with the file in Migrating status.

Here's an example showing immediate return and changing state:

```
% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 12:43 (REG) C

% dmput C ; dmls -l
[immediately:]
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 12:43 (MIG) C

[then after a while:]
% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 12:43 (DUL) C
```

Example showing delayed return, and release of disk space:

```
% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (REG) C

% dmput -r -w C ; dmls -l
[long pause:]
total 220
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (OFL) C
```

The *dmfind* command

The *dmfind* command is based on the GNU version of the find command and adds the ability to search for files in a given DMF state. This is handy for determining which files are offline and, therefore, candidates for retrieval with *dmget*.

Example:

```
% dmls -l
total 220
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (OFL) C

% dmfind  .  -state ofl  -print
```

```
./C

% dmget  `dmfind . -state ofl -print`
[pause:]

% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (DUL) C
```

To efficiently recall all files in a directory named mydir and its subdirectories, use the following command:

```
% dmfind mydir -state mig -or -state ofl -print | dmget
```

The *dmcopy* command

The *dmcopy* command copies all or part of an offline file to a destination file, keeping the offline file offline.

When the offline file is being copied in its entirety, the only arguments needed are the two filenames.

Example:

```
% dmls -l
total 224
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (OFL) C

% dmcopy  C  newC
[pause:}

% dmls -l
total 1404
-rw-r-----   1 aeneuman madmag      20155 Mar  2 11:24 (REG) A
-rw-r-----   1 aeneuman madmag     201550 Mar  2 11:24 (REG) B
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (OFL) C
-rw-------   1 aeneuman madmag    1209300 Mar  4 15:06 (REG) newC
```

Dmcopy has options that allow copying just part of the offline file, and specifying offset locations in the source and destination files.

-l length	Specifies a data length to copy in bytes. The default is the whole size of the source file.
-o source-offset	Specifies a byte offset in the offline source file where reading is to begin. The default is zero; that is, reading begins at the start

	of the source file.
-d destination-offset	Specifies a byte offset in the destination file where writing is to begin. Any bytes in the destination file before this offset are zero filled. The default destination offset is whatever the source offset is.

Example: Skip two records of the offline source file and copy the next three records. Records are 2K bytes long.

```
% set RECDLEN=2048
% set NRECD=3
% @ LENGTH = $NRECD * $RECDLEN
% set NSKIP=2
% @ OFFSET = $NSKIP * $RECDLEN

% dmcopy  -l $LENGTH  -o $OFFSET  -d 0  C  newC
[pause:]

% dmls -l *C
-rw-r-----   1 aeneuman madmag    1209300 Mar  3 15:17 (OFL) C
-rw-------   1 aeneuman madmag       6144 Mar  4 16:10 (REG) newC
```

Note: If the source file is a regular file, without an offline copy in DMF, the destination file is always zero length.

Note: The modes on the source file are ignored when creating the destination file. The modes on the destination file are 0666 (readable and writable by everyone), except where blocked by the current umask. Review your output files' permissions and reset them with *chmod* as needed.

The *dmattr* command

The *dmattr* command prints selected attributes of specified files. This is most useful in shell scripts.

Some options include:

-a attr1,attr2,...	Selects a subset of the reportable attributes.
-d delimiter	Specifies a one-character separator between adjacent values. A space is the default.
-l	Prints the attributes, one per line, with labels.

Examples:

```
% dmattr -l A
      bfid : 0
     emask : 0
```

```
   fhandle : 01000000000000188dede3a4b319c5b9000e00000000000100000000b3fd163
     flags : 0
     nregn : 0
     owner : 4771
      path : A
      size : 20155
     space : 20480
     state : REG

% dmattr A
0 0 01000000000000188dede3a4b319c5b9000e00000000000100000000b3fd163 0 0
4771 A 20155 20480 REG

% dmattr  -a owner,state  -d :  A
4771:REG

% foreach file ( * )
? if (`dmattr -a state $file` == OFL) then
? echo $file is offline
? endif
? end

C is offline
```

Storage Best Practices

Portable File Names and Sizes

DRAFT

This article is being reviewed for completeness and technical accuracy.

Portable File Names

Use portable file names. A name is portable if it contains only ASCII letters and digits, \`.', \`_', and \`-'. Do not use spaces or wildcard characters or start with a \`-' or \`//', or contain \`/-'. Avoid deep directory nesting.

If you intend to have tar archives to be read under MSDOS, you should not rely on case distinction for file names, and you might use the GNU **doschk** program for helping you further diagnose illegal MSDOS names, which are even more limited than Unix like operating system.

Portable File Sizes

Even though Lou's archive filesystem will allow a file size to be greater than several hundred gigabytes, not all operating systems or filesystems can manage this much data in a single file. If you plan to transfer files to an old Mac or PC desktop you may want to verify the maximum filesize it will support. Likely a single file will need to be less than 4 GB before it will transfer successfully.

Dealing with Slow File Retrieval

DRAFT

This article is being reviewed for completeness and technical accuracy.

There are sometimes problems with commands on Lou, that should finish quickly, but end up taking a long time.

When you do an " **ls** " on Lou, you see all the files on disk that you've put there. However, most of the files are actually written to tape using SGI's Data Migration Facility (DMF).

One problem with DMF is that it does not deal well with retrieving one file at a time from a long list of files. If you do an " **scp** " with a list of files, Unix feeds those files to DMF one at a time. This means that the tape(s) containing the files is getting constantly loaded and unloaded which is bad for the tape and tape drive, and also very slow. As the list of files gets longer (by use of "*" or moving a "tree" of files) the problem grows to where it can take hours to transfer a set of files that would only take a few minutes if they were on disk. When several people do file transfers at once that retrieve files one at a time, it can tie the system in knots.

Optimizing File Retrieval

DMF let you fetch files to disk as a group with the **dmget** command. The tape is read once and gets all the requested files in a single pass. Essentially, give **dmget** the same list of files you are about to transfer, and when the **dmget** completes, then **scp**/**ftp**/**cp** the files as you had originally intended. Or you can put the **dmget** in the background and run your transfer while **dmget** is working. If any files are already on disk, dmget sees this and doesn't try to get them from tape.

There is also a **dmfind** command that let you walk a file tree to find offline files to give to **dmget** . Make very sure you are in the correct directory before running **dmfind** . Use the " pwd " command to determine your current directory.

Please check to make sure too much data isn't brought back online at once by using du with the *--apparent-size* option or by using **/usr/local/bin/dmfdu** .

Note that **dmfdu** will give an error message for each symbolic link that points at a nonexistent file.

```
lou% /usr/local/bin/dmfdu Foo
    Foo
             13 MB regular           340 files
           1114 MB dual-state       1920 files
          74633 MB offline          2833 files
             13 MB small             340 files
```

```
        75761 MB total           5093 files
```

When transferring data between Lou and Columbia nodes use the /nobackup filesystems, instead of the Columbia NFS (slow) home directories.

File transfer rates vary depending on the load on the system and how many users are transferring files at the same time. Transferring files using **scp** between Lou and Columbia nodes on the /nobackup file system for files larger than 100 megabytes is typically between 7 - 17 MB/s using the gigabit network interface. Transferring files using scp between the Columbia nodes for files larger than 100 megabytes, is typically between 20-30 MB/s for the gigabit network interface.

Example 1:

```
lou%  dmget *.data &
lou%  scp -qp *.data  myhost.jpl.nasa.gov:/home/user/wherever
```

Example 2:

```
lou%  dmfind  /u/username/FY2000 -state OFL -print | dmget &
lou%  scp -rqp /u/username/FY2000 some_host:/nobackup/user1/whereever
```

You can see the state of a file by doing " **dmls -l**" instead of " **ls -l**". For more information on using DMF, please look at:

Data-Migration-Facility-DMF-Commands_250.html

Maximum Amount of Data to Retrieve Online

The online disk space for Lou1-3 is much, much less than its tape storage capacity, and it is impossible to retrieve all files to online storage at the same time. So, before retrieving a large amount of data, you should check that there is enough online space for it. The **df** command shows the amount of free space in a filesystem. The Lou script **dmfdu** reports how much total (online and offline) data is in a directory. To use this script, simply " **cd** " into the directory you want to know total amount of data for all the files in the current directory and execute the script.

If you would like to know the total amount of data under your home directory on Lou, you need to first find out if your account is under s1a-s1e, s2a-s2e or s3a-s3e. Assuming you are under s1b, you can then use **dmfdu /s1b/your_userid** to find the total amount. Another alternative is to simply cd to your home directory and use " **dmfdu *** ", which will show use for each file or directory.

Lou1-3's archive filesystems are between 8 TB and 30 TB in size, but the available space typically floats between 10% to 30%. In the example 3, 29% of space is unused. It is best to retrieve at most 10% of the filesystem space at a time. Do what you need to with those files (scp, edit, compile, etc), then release (dmput -r) the space, and then retrieve the next group of files, use them, then release the space, etc. For example 3, retrieve one directory's data from tape, copy the data to remote host then release the data blocks, before retrieving more data from tape.

Example 3:

```
lou% df -lh .
Filesystem            Size  Used Avail Use% Mounted on
/dev/lxvm/lsi_s1b     8.6T  6.1T  2.6T  71% /s1b



lou% dmfdu project1 project2
project1
         2 MB regular            214 files
        13 MB dual-state           1 files
    229603 MB offline            101 files
         2 MB small              214 files
    229606 MB total              315 files

project2
         7 MB regular            245 files
      4661 MB dual-state          32 files
    218999 MB offline             59 files
         7 MB small              245 files
    223668 MB total              336 files


lou% cd project1

lou% dmfind . -state OFL -print | dmget &

lou% scp -rp /u/username/project1 remote_host:/nobackup/username

(Verify that the data has successfully transferred)

lou% dmfind . -state DUL -print | dmput -rw

lou% df -lh .


lou% cd ../project2

lou% dmfind . -state OFL -print | dmget &

lou% scp -rpq /u/username/project2 remote_host:/nobackp/username

lou% dmfind . -state DUL -print  | dmput -rw
```